Proceedings
of the International Workshop
"Innovation Information Technologies:
Theory and Practice"

Dresden, Germany

September 06-10, 2010

Proceedings of the International Workshop
"Innovation Information Technologies: Theory and Practice"
Dresden, Germany, September 06-10, 2010

ISBN: 978-3-941405-10-3

This International Workshop is a high quality seminar providing a forum for the exchange of scientific achievements between research communities of different universities and research institutes in the area of innovation information technologies. It is a continuation of the Russian-German Workshops that have been organized by the universities in Dresden, Karlsruhe and Ufa before.

International Workshop "Innovation Information Technologies:
Theory and Practice"

Organized by
Forschungszentrum Dresden-Rossendorf (FZD), Dresden, Germany
Ufa State Aviation University (USATU), Ufa, Russia
Karlsruhe University of Technology (KIT), Karlsruhe, Germany

The workshop is supported by
German Research Association (DFG)
Forschungszentrum Dresden-Rossendorf (FZD)

General Chairman: Peter Joehnk, FZD, Dresden, Germany
General Chairwoman: Nafisa Yusupova, USATU, Ufa, Russia

Editors: Uwe Konrad, FZD, Dresden, Germany
Liliya Iskhakova, USATU, Ufa, Russia

Published by Forschungszentrum Dresden - Rossendorf

Available from Forschungszentrum Dresden - Rossendorf
Library Services
Bautzner Landstr. 400
01328 Dresden
Germany
e-mail: contact@fzd.de
http://www.fzd.de

ISBN 978-3-941405-10-3
Open Access http://nbn-resolving.de/urn:nbn:de:bsz:d120-qucosa-39940

Contents

Modern Trends in the Development of Software Applications

Daniel Petrik
MMS SOFTEC GmbH
Trnava, Slovakia
e-mail: petrik@mms-softec.sk

Oliver Moravcik
Institute of Applied Informatics and Automation
Slovak University of technology
Trnava, Slovakia
e-mail: oliver.moravcik@stuba.sk

Abstract[1]

The pros and cons of using an object relational database mapper when compared with direct SQL commands. Separation of the user interface and of the logic from a business application. Reasons for the separation or for free linkage between the user interface (UI) and the application.
Model View Controller (MVC) and the Model View Presenter (MVP), patterns for application development. Description of individual parts MVP, their place in the application, composition of user interface (UI) form modules and the links between modules.
UI composition from visual objects, graphic elements, templates and styles. A comparison of options offered by a visual composition of UI objects, clarifying the concept of "lookless" graphic visuals and defining their appearance. Automatic UI layout adjustment, scaling, animations etc. An example with WPF controls.
Linking visual elements (controls) with visualised data – data binding and automatic updates. Comparison of the standard method of connecting the visual elements to visualised data with data binding applied in WPF, an explanation of automatic updates (one / two-way) about the content of visual objects and the binding data objects.
Managed languages – C# - ways of diagnosing the possibility of a memory leak and it's origination and causes in the managed application environment
Test Driven Development (TDD) – a description of some problems from the developing praxis which can complicate and set limits by TDD.

1. OR Mapper (Object-relational database mapper).

The main reason for the use of the OR Mapper is to shield the software developer from any relation to the database. The developer can fully focus on working with objects and their collection whilst not having to bother about SQL commands and their relational database schema. After a few tests and reading other test reports we decide to use the OR Mapper (http://www.devexpress.com/Products/NET/ORM/).
Why we decided to use this OR Mapper:
It has an ability to create database schema from class definition and is able to update database schema when there are modified classes. This is very important while we generate these classes from the UML class model.
Supported databases MS SQL, Oracle.
Automatic mapping relations between objects (1:1, 1:N, M:N) to the database structure.
It has the ability to use LINQ expression to query objects, support for filtering and transactions.
The OR Mapper in our case is responsible for all CRUD operations, there are no direct link to a related database.
All classes implement automatic property change notification, which allows automatic updates of UI controls by using this object.

The main limitation (as we discovered in the development process):
There was a problem with processing large amounts of data – modification of any object's attributes require the reading of the object in memory. This can lower performance, for example when it is a deleted root object with many child objects or with a more deeper hierarchy (objects in tree), all objects to delete should be first read from the database in memory and then deleted. We solved this problem in the most critical part using a little hack – defining cascade delete in the related database.
The impossibility to call a database procedure or function - In some cases we used a calculated property, this requires that the whole object is read from the database. This is a limitation when we query objects stored using a calculated property as a part of the condition for querying.
The impossibility to join objects in a query - when there is no relation between objects – a workaround is to define views on the database and then query the defined views as an object.

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

2. Modern UI composition

Actual "rich" business applications typically feature multiple screens, rich, flexible user interaction, data visualization, and role-determined behaviour. The application's expected lifetime is measured in years and that it will change in response to new, unforeseen requirements. This application may start as small and over time evolve into a composite client. Composite applications use loosely coupled, independently evolvable pieces that work together in the overall application. This is a path for modularity and the application of patterns. The base of modular (composite) application is to combine individual views into a composite view this often being the main (composing) view (shell) defined layout for the child views. For management of coupling views with data the following patterns are used: Model-View-Controller, Model-View-Presenter or the Model-View-Presentation Model. More info about patterns: http://martinfowler.com/eaaDev/uiArchs.html, http://en.wikipedia.org/wiki/Model%E2%80%93view%E2%80%93controller .

When studying patterns for the view composition and the library Prism application, we found out that we implement the MVP pattern. As development continues we see that the implemented pattern MVP is closer to the Model-View-Presentation Model pattern. This is because we use a model for the database manipulation and operation via the use of the OR Mapper together with the inclusion of valid business rules while the view (WPF control) is coupled with the Presentation Model. The Presentation Model does required data transformation for Viewing, manages responses on user events and handles transition states, i.e. data editing and inserting.

To keep the Presentation Model as simple as possible, we decide to use "in view", constructing a UI composition technique like templating, styling, and data binding with automatic change notification. Because of the requirements of the Windows platform, we were forced to use Window Presentation Foundation (WPF) which is where the term "lookless" control comes from. It is necessary to explain that the WPF visual control has no look, while the look of the control is defined by their templates and styles. The control template defines base UI visualization, e.g., a button can be defined as a rounded rectangle. When control needs to show raw data, i.e. some object, it needs a data template to know how to visually represent the data. The connection of raw data with a data template is a used mechanism called data binding. Data binding can be unidirectional or bidirectional. As a result the visual component can automatically update itself as the bound data is changed (e.g., objects properties), and data can change as a result of user action. Data binding can use converters to convert data to the required representation. The following picture shows how the visual output of visual control is composed.

The rendered output consist of raw data (in this case object Address), from template(s) defining positions of UI items (label, textblock, image), from style and layout (in this case the style defined is that of ZIP and the city value should be rendered with bold font and from data binding, which binds the property of data object with UI items (in this case textblocks, image) which can also call a converter to convert the Map data to a picture.

Another feature of WPF control is layout. Layout provides sizing and positioning. Layout, as implemented, allows the creation of resolution independent of UI, and creates control with automatic size measurement to correctly show its content.

Styles – styles have the same purpose as CSS for web applications. Style provides an abstraction that can gather up all properties of an element. For example when using style at an application level we can define the property of the appropriate elements in all of the application. Style can also define (redefine) data and control templates, thus by style it is possible to totally redefine how a visual component is rendered on UI. Style can also define triggers, to modify a visual look

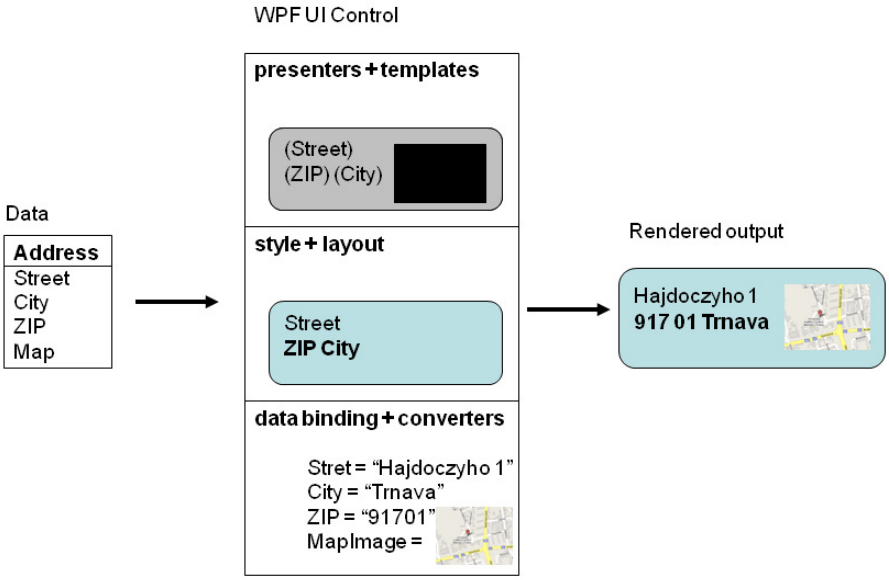


Figure 1. Visual output composition

as a result of some data or state of change.
Events – we present a few words about events.

As shown in Figure 1, visual output is rendered from data templates and styles. Data binding as is implemented in WPF allows the automatic re-rendering of a visualization of data when the data property notifies visual control about its state of change. In comparison to WinForm, it allows the developer to "forget" about manual refreshing which is usually realized as a method of setting actual values to visual components.
Data binding can be realized as a two-way process, when change in a data object is propagated to visual control and data change in visual control (as a result of user action) is propagated back to the data object, one-way, from the data object to the visual object, or from the visual object to the data object.
I have to note that it is also possible to define triggers (in style and template), these triggers can create a modification of the visual control properties. For example, an automatic background colour change when a value is missing or out of range.

3. Managed languages – C# - possibility of the memory leak origination

The managed application of memory management is managed by an application framework called a "garbage collector" (GC). There are many situations when dependency between objects does not allow the GC to dispose of objects from memory, this was discovered when testing the WPF application using the MVC pattern. We also found out that some presenters exist in memory more than once. Analysis of this situation shows this view, the WPF composite user control keeps reference to data objects published by the presenter. The solution was to implement and "INotifyPropertyChanged" in the interface with the objects data otherwise the "DependancyProperties" class and the "INotifyPropertyChanged" interface are unavailable.
WPF uses the ValueChanged event, which involves calling the "PropertyDescriptor.AddValueChanged" method than the "PropertyDescriptor" object that corresponds to property. Unfortunately, this action causes the common language runtime (CLR) to create a strong reference from this "PropertyDescriptor" object to the data object. The CLR also keeps a reference to the "PropertyDescriptor" object in a global table (see http://support.microsoft.com/kb/938416/sk). Another source of memory leaks is in the use of command binding. This was the case when we started to use Prism 1. The solution was to extend presenters by the method to "Unsubscribe", when the view (window) is going to close in order to remove references to commands. In the next version (2.0) of the prism (called "Composite Application Guidance for WPF and Silverlight"), this was solved using weak references.

The diagnostics of memory leaking seems to be simple, especially when a memory profiler is used. This tool can help to find out "paths to a GC root", which can be used to identify a problem object and its references. Note, when profiling the memory of .NET 4 – an object can be kept in memory until there is a new demand for free memory.

4. Conclusion

Modern trends in an application software development leads to programmer friendly development process, which eliminates routine tasks. This allows development teams to shorten time necessary to build high added value software product with respect to user experiences.

References

1. http://blogs.msdn.com/b/jgoldb/archive/2008/02/04/finding-memory-leaks-in-wpf-based-applications.aspx
2. http://support.microsoft.com/kb/938416/sk
3. http://www.devexpress.com/Products/NET/ORM/.

Internet-based multimedia/multimodal User Interfaces for Teleoperating Robots (ImRoNet)

Prof. Dr.-Ing. Heinz Wörn Dipl.-Inform. Michael Mende Dipl.-Inform. Simon Notheis
Institute for Process Control Automation and Robotics (IPR)
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
e-mail: heinz.woern@kit.edu michael.mende@kit.edu simon.notheis@kit.edu

Abstract

This paper presents the ImRoNet project, which aims at providing new Internet-based user interfaces for teleoperating robots by using multimedia and multimodal technologies. A modular architecture is introduced that allows for reliable and secure transport of data between the different software and hardware modules. Also, miscellaneous components necessary for the interaction are presented in detail. Furthermore, two scenarios involving local and remote support for a worker are described.

1. Introduction

When looking into the current development in the field of service and industrial robotics, one can see a paradigm shift on several levels. The rigid industrial robots are becoming flexible and networked robot systems, the boundaries between concepts of service robotics and the concepts of industrial robotics become blurred, and more and more tasks require a high degree of interactivity between human and machine. Novel user interfaces allow a new form of interaction with the robot and open up new fields of applications for robots. Several preconditions must be met to achieve the goal of an intuitive way to control robots. We need flexible multimedia-based and multimodal interfaces, e.g. of visual, aural or haptic type, some of them are already known from other areas of human-machine-interaction We need flexible and adaptive robot systems that provide (semi-)autonomous functions and skills with a high degree of individual intelligence. Furthermore, local as well as remote control has to be provided.

The primary research goals of ImRoNet are related to the development of new methods of multimedia / multimodal interaction, the development of new operating and control devices, the transfer of information over the Internet as well as the development of autonomous robot functions to provide user support. Different semi-autonomous control systems are developed and realized that enable a remote user to affect the events on site by means of a robot or to assist a local user as an expert.

Possible applications for such teleoperation systems are the supervision and control of technical facilities, buildings, as well as their surroundings. In essence, three modes of operation with multimodal interfaces can be defined: Tele-assistance, tele-presence and an autonomous assistance/information system.

Project partners beside the IPR are Fraunhofer IPA, BECKHOFF Automation GmbH, metaio GmbH, GPS GmbH/Neobotix, and SCHUNK GmbH & Co. KG.

2. System architecture

To support local as well as distributed software components, the ImRoNet architecture is based on CORBA (ACE/TAO) as underlying middleware. All software components in ImRoNet share a common, well-defined interface and data flow structure. The realized components can be divided into 12 major blocks as shown in Fig. 1. A typical data flow may e.g. contain an initial user input via some kind of input device or GUI, input data preprocessing or abstraction, derivation of the intended command, validation, planning and execution of the derived command, and appropriate feedback provided to the user. The modality used for user feedback may depend on the user's preferences or the surrounding. The Modality Renderer may transform e.g. distance information into different kinds of output signals (like sound, light, vibration, …).

Due to the well-defined interfaces and data structures, as well as the different data abstraction layers in the architecture, components can be easily added or replaced, according to changes in the hardware setup or algorithms used in the scenario. Furthermore, components can be easily configured and reconnected graphically during runtime to allow switching between different scenarios reusing the existing modules.

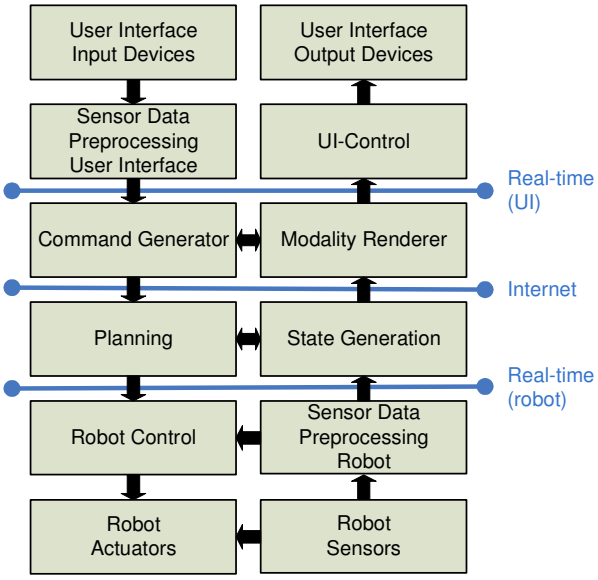


Figure 1. The ImRoNet architecture

3. Components of interaction

In this paragraph, we provide a short overview of the hardware and software components used and developed to provide the necessary interfaces. These include the discussion of various input and output devices like projectors and head mounted displays as well as the components necessary to develop a novel control panel.

3.1. Projector based interaction

In general, several options can be distinguished, depending on the mounting point of the projector. Since the main focus lies on mobile interaction, fixed-mounted projectors and those mounted on a kinematic chain are less interesting. Instead we concentrate on two possible mobile scenarios: a projector mounted on the robot, in our case on a pan/tilt unit beside the manipulator, and a head mounted miniature projector on a helmet to be worn by a worker (see Fig. 2). In both cases, the primary use case is supposed to be local support by augmenting the environment or providing flexible interaction interfaces.

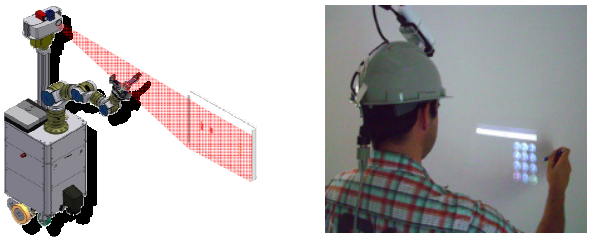


Figure 2. Projector based interaction

The tasks required to provide projector based interaction can be divided into three classes. First of all, calibration methods for the projector and camera are necessary, that have to be extended to allow projector-camera-system (PCS) calibration and the combined calibration of PCS and input device, e.g. a pen.

Secondly, challenges concerning visualization have to be overcome, such as rectification of the image plane and rectification of the displayed image on the surface. To accomplish rectification on (almost) arbitrary surfaces, surface reconstruction by means of structured light (e.g. stripe projection) is self-evident, because all needed parts are already present in the given hardware setup. Thirdly, the interaction methods themselves are developed. This ranges from rather simple tasks like on screen keyboards or drawing to complex interaction with the system.

3.2. HMD based interaction

Head mounted displays (HMDs) offer an easy and flexible solution to interact with the environment. The wearer can either be a local worker or a remote expert. In the first case, to ease the overlay of the camera image and the augmentation, only video-see-trough devices are used, so the HMD is equipped with a camera to capture the working environment. In the latter case, no perception of the local (expert's) environment is necessary at all, but the user should be provided with an intuitive and immersive impression regarding the robot's environment. The HMD is furthermore equipped with an inertial sensor to track the head movement. To provide input commands, the remote user uses a 6-D mouse or similar devices.

3.3. Control panel

We developed a platform to evaluate and implement multimedia-based and multimodal user interaction scenarios (see Fig. 3, left). The modular hardware setup serves as a testing platform for new ways of interaction with mobile and industrial robots, as well as for interaction with the robot's environment and/or simulations including models used e.g. for path planning tasks. Usability and ergonomics in common control panel designs are improved by adding new intuitive modalities like multi-touch screens, inertial sensors, together with haptic and visual feedback rarely used in commercial robot control devices so far.

The hardware is based on a low-cost, netbook with a resistive touch screen. Miscellaneous input components were additionally included, such as a 6-DoF inertial sensor, a compass and a 6-DoF multi axis controller. Image acquisition is provided by a forward directed main camera for scene/cell monitoring and a user camera. Although the main use of the user camera is supposed to be the communication via video chat with a remote expert, eye tracking algorithms were

included to evaluate its UI control possibilities, or to provide an attention analysis.
Multimodal output can be generated via the built-in speakers, a vibration module with vibration motors, RGB LED modules to illuminate the panel and red/green laser diodes to interact with a projector-camera-system or to mark points seen by the forward scene camera..

Regarding the vibration module, various ways of haptic output while interacting with the GUI depending on wave forms and the quantity and type of vibration generators were evaluated. Furthermore, algorithms for the inertial sensor and compass were implemented that incorporate filtering, sensor data fusion and robustness, to allow for inertial 6-D robot control (either full strap-down 6-DoF or two-step, i.e. rotation or translation consecutively), optionally with haptic feedback.

One key component of the software architecture is the Modality Renderer, that implements, among other things, the automatic mapping of inputs or misc. sensor data (e.g. distance to obstacles or collision warning) to output modalities (sound, light, vibration, …) based on environmental conditions or user preferences. The multi-touch capable GUI provides easy 2-D platform control and 6-D arm control and served to evaluate intuitive ways of interaction in combination with the attached sensors and output devices.

Some of the features evaluated on the development panel were also transferred to a novel panel design by Beckhoff which is designed for commercial use in industrial environments (see Fig. 3, right).

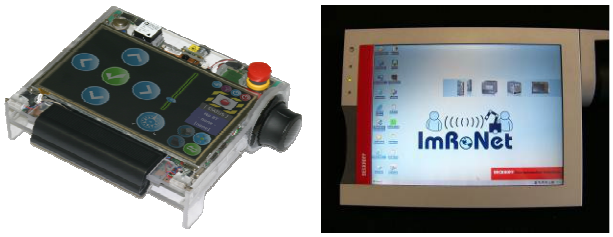


Figure 3. Development panel (l), Beckhoff panel (r)

3.4. Platform, Arm and Gripper

The robot systems consist of mobile platforms by Neobotix, a Schunk LBR 7-axis arm and a Schunk Dextrous Hand (SDH) which is a versatile industrial 7-DoF three-finger-gripper equipped with tactile sensors for contact imaging. An RFID reader has been integrated into the SDH to facilitate the selection of the best grasp skill for the particular object based on Internet database information. Furthermore, an electronics extension provides full module maintenance via Bluetooth in the field and independent of the CAN bus. Thus real-time observation of module communication during operation is possible.

4. Scenarios

In the following sections we will present local and remote user scenarios based on two different robot platforms. The local or remote user is supported by means of augmented reality features, either overlaid onto a camera image or projected directly into the environment using a projector.

4.1. Local maintenance support

This scenario is based on a mobile platform equipped with a 7-axis arm and a panel mounted on its end effector, whereby the panel features a 6-D controller (Fig. 4, left). The worker grasps the controller and is able to guide the robot by dragging it in the desired direction. After arriving at the object that needs maintenance, e.g. a control cabinet, the panel can be moved in front of the object. Automatic robot localization or model-based tracking allows for computing the current position of the panel's built-in camera w.r.t. the target object. This provides the necessary information to overlay working instructions by means of augmented reality (Fig. 4, right). Another application is the transmission of process parameters (e.g. rpm) via industrial standard protocols like OPC/UA, which can be augmented into the camera image as well.

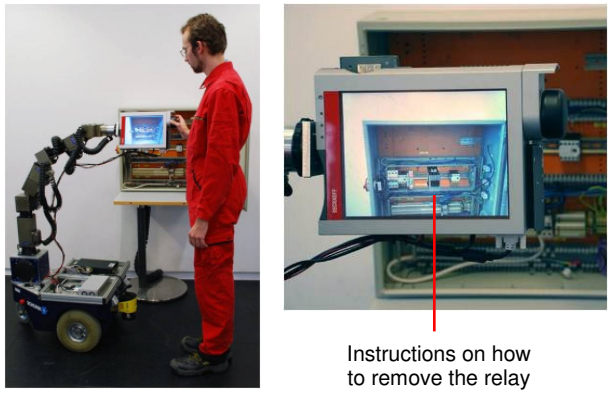


Figure 4. Demo 1: Local maintenance support

In a second example, a robot mounted projector displays GUIs or information about hidden structures like power lines or pipes within a wall directly onto the corresponding place on the surface and thereby also assist local workers in their tasks (see Fig. 5).

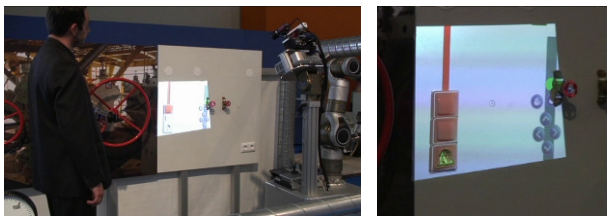


Figure 5. Demo 2: Local maintenance support

4.2. Semi-autonomous maintenance and inspection

The second scenario demonstrates teleoperated movement of a robot in a remote environment. The robot is controlled via a haptic input device (Falcon), while the virtual representation of the remote environment is displayed on screen (see Fig. 6). To reduce the danger of damaging the robot or equipment and to ease the control, collision avoidance is provided by the robot.

An automatic detection of handwheels by the robot's 3-D ToF sensor system takes place, followed by highlighting the handwheels on the screen. The manipulation itself is semi-autonomous, so the robot can react promptly e.g. based on tactile or force/torque sensor data from the gripper. While selecting the handwheel, the user is supported by haptics through a magnetic effect to make the approach to the handwheel easier, while the robot moves to an appropriate position, grasps and rotates the handwheel in the desired direction. To operate large handwheels, a coordinated movement of platform and arm occurs.

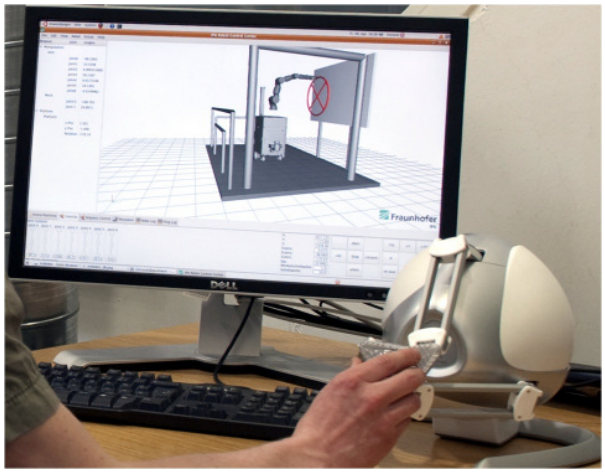


Figure 6. Virtual scene view and haptic input device

4.3. Immersive direct teleoperation

In this scenario, an actual camera image (robot mounted camera or scene overview camera mounted above the robot cell) is augmented with GUI elements and extra information to facilitate the robot control. The remote user can switch between different control and view modes and thereby control the robot platform via a 6-D mouse or control the camera movement by mapping head movements captured by the HMD's built-in inertial sensor to the pan/tilt units carrying the cameras. All kinds of input commands resulting in robot or camera movements are transformed into their corresponding absolute target positions and rendered into the scene model and camera image. That way, the user has direct feedback about his commands which compensates the latency in the control loop between the (remote) user and robot, i.e. the latency in receiving the robot's current pose and/or the camera's video stream.

Furthermore, data derived from the parallel simulation based on the 3-D model of the scene and robot such as distances, virtual representations of the environment and current view/movement directions can be augmented into the camera image as well. Fig. 7 shows the user's view seen from a camera mounted on a pan/tilt unit above the cell. The user can control the camera and robot via head movements or other input devices and is presented with additional information rendered into the camera image.

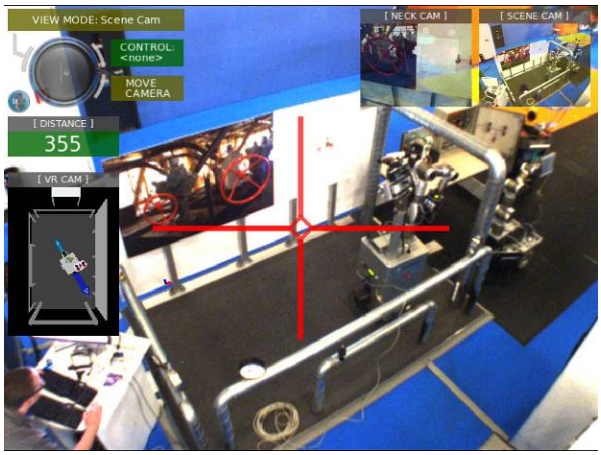


Figure 7. The remote user's view through a camera mounted above the scene, including augmented information such as distances and additional virtual and real camera views rendered into the image.

5. Conclusion

Within the ImRoNet project, flexible and modular software and hardware components were developed and new ways of multimedia-based multimodal interaction were evaluated and realized. Those components and concepts were integrated into two common demonstrators and presented in June 2010 at the AUTOMATICA fair in Munich as part of the final project presentation.

Acknowledgements

Special thanks go to the BMWi and DLR for supporting and funding this project and to all our partners within the project.

References

1. ImRoNet Homepage
   http://www.imronet.de
2. DLR project page
   http://www.pt-multimedia.de/de/97.php

Innovation Information Technology in the Service of Research at the FZD

Uwe Konrad
Forschungszentrum Dresden-Rossendorf e.V.
Department of Information Technology
Dresden, Germany
e-mail: u.konrad@fzd.de

Abstract[1]

Large research facilities, such as the Forschungszentrum Dresden-Rossendorf (FZD), provide the scientific work environment for employees, guests and students from all over the world.

To manage research projects, achieve new insights and visualize the results innovative information technology has to be applied for the communication, information and collaboration infrastructure as well as for the scientific equipments. Managing such a dynamic environment in a networked world is a challenging task. This paper describes some solutions and implementation aspects of innovative information technology for a large scale research facility developed in Dresden. The solutions are illustrated using some examples out of the areas of High Performance Computing (HPC), Virtual Reality (VR) and Collaboration in Virtual Institutes.

1. Introduction

The FZD is engaged in basic and application-oriented research. The main topics are Advanced Materials Research, Cancer Research and Nuclear Safety Research. Six major research facilities are operated by the FZD, which are also available to external users:
Radiation Source and Free Electron Laser (ELBE), Ion Beam Center, Rossendorf Beamline at the European Synchrotron Radiation Facility (ESRF), Positron Emission Tomography Center, TOPFLOW (a multipurpose thermo-hydraulic test facility) and the Dresden High Magnetic Field Laboratory.
Currently more than one thousand people are active on the campus.

2. Concepts and Solutions

2.1. High Performance Computing

The demand for computing power and storage capacity in the area of research is growing fast, so at the FZD. The key questions to meet the future demand of computing power are on three areas: energy consumption, bandwidth scaling and programmability.

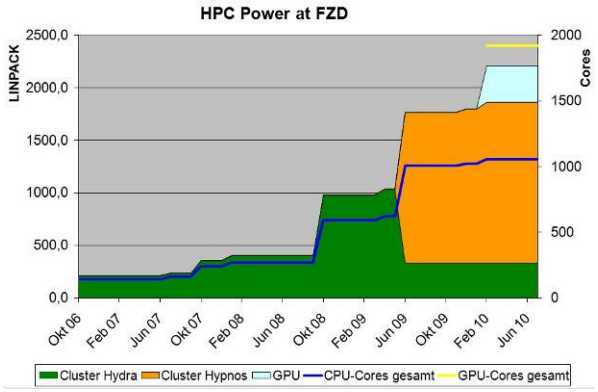


Figure 1. HPC at the FZD

Multi-core systems are still the standard solution; many-core processors (chips with more than 20 up to hundreds of processors) are on the horizon; and many-core boards such as the NVIDIA GPUs are available.

Unfortunately, manufacturers cannot add too many cores to their chips because of bandwidth limitations both in moving data to and from external memory as well as amongst the on-chip processing cores. The additional cores quickly become ineffective because they are starved for data.

Modern multi-threaded algorithms and software for many-core architectures may need to adapt to better exploit locality and streaming behavior. This is a similar challenge as transition to clusters and MPI in the past.

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

Modern computing clusters combine multi-core central processing units (CPU) and many-core graphics processing units (GPU). Figure 1 shows the development of the FZD cluster during the last years indicating the introduction of GPUs recently.

GPUs have been originally developed to perform computation for computer graphics. Using the CUDA software development kit Nvidia GPUs can be used for general-purpose computation, too. Despite many improvements to GPUs the functionality is still limited. Especially memory operations are a bottleneck due to the limited bandwidth of the bus systems.

The implementation of an approach for fast parallel computation in the field of statistical physics at the FZD [1] shows the great potential of GPU-based solutions. Other processor manufacturers such as AMD and soon Intel also deliver GPU for general purpose computations. Therefore vendor independent development kits are needed for the combined CPU and GPU parallel computing tasks. The Open Computing Language (OpenCL) with the OpenCL C programming language was first introduced by Apple in 2008 and later standardized by the main processor vendors. This standard may serve as a more platform independent development kit for future CPU and GPU programming.

2.2. Virtual Reality and 3D Visualization

In the last years Virtual Reality and 3D Visualization have come to a more realistic exposition. This innovative technology has found its way not only into the entertainment or manufacturing industries but also to science and research. But it also needs massive computing power to produce such complex visualizations. For example, to render James Cameron's movie Avatar (released end of 2009), a server farm with 35,000 CPU cores was used.

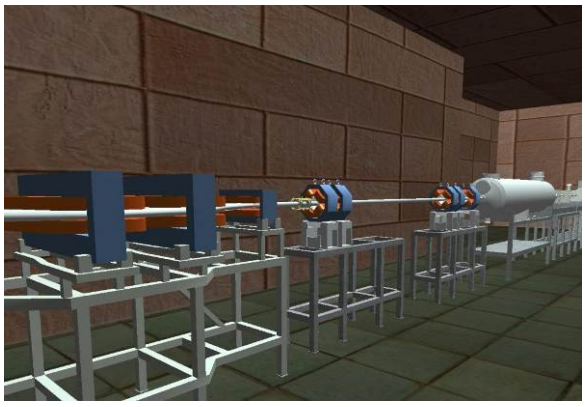


Figure 2. VR Simulator "ELBE" at the FZD

Also at the research center in Dresden VR implementations are used. Together with the Technical University Dresden, Faculty of Computer Science a simulator which employs VR techniques to implement a test and training environment has been developed [2]. It provides a computer simulation of the superconducting linear accelerator for electrons (ELBE).

With this software you can not only control the electron beam but it is also possible to make a virtual tour through this giant device and see how the main components work. The next step of this development is an implementation for the new Centre for High-Energy Radiation Sources currently built at the FZD.

Basis for a VR Simulation is a detailed model of the physical system but in addition the behavior of the users and the solutions for input and output systems are of special interest. This is a major challenge for the second example of a VR implementation at the FZD: a fire fighting simulation software [3]. It provides different scenarios for firefighting, rescue and search operations and offers visual, audio and haptic feedback to the users. Special care has been exercised to model the fire with its energetic behavior the water and foam ejection and the superposition of radiation sources.

Figure 3. Firefighting Simulator

The system has been developed and tested with the professional campus fire brigade at the FZD showing proper results. Trainees at FZD participate in the development process in order to strengthen their skills.

2.3. Collaboration and "Virtual Institutes"

Information Technology is not only the foundation for the scientific and technological projects but also the basis for the collaboration of the people involved. Therefore it also needs innovative forms of organization and communication. Such an approach is the organization within Virtual Institutes.

Virtual Institutes enable groups of people from different geographical locations to organize and meet together for presentations and exchanges of various

kinds. At the German Helmholtz research Association Virtual Institutes are in practice since a few years. 87 Virtual Institutes in which almost 217 university partners from 55 German universities are actively involved. The FZD is also involved in some virtual institutes.

To implement a virtual institute state of the art communication tools and technologies are used. This includes multipoint videoconferences, browser-based real-time meeting applications (such as Acrobat Connect, WebEx), web-based collaborative workspace (e.g. NX, SVN), GRID communities or team portals (SharePoint et al.).

To provide scientists a suitable infrastructure the German Priority Initiative "Digital Information" was founded by the alliance of the German science organizations. The area of virtual research environments is one of the 6 fields of activities. Since 2010 there is also a working group within the European Knowledge Exchange network engaged in this field.

A challenge in such a dynamic environment is the Identity Management (IdM) of all the users involved. Not only the individual identities, but also the roles and groups of identities have to be managed together with other resources. The identities are embedded in local and global workflows. The solution for this task on the corporate level at the FZD was described before [4]. But for a virtual research environment a more global Authentication and Authorization Infrastructure (AAI) is required.

An AAI federation creates an organizational framework and a trust relationship between federation members (facilities and providers). For the science and education community in Germany a central instance of an Authentication and Authorization Infrastructure is the DFN-AAI. The implementation is based on the open source software Shibboleth (developed by the INTERNET2-Initiative) and the use of digital certificates (X.509).

Over 100 universities and research institutes are already member of this federation, but also large international publishing companies as well as Microsoft Germany.

3. Conclusion

Innovation Information Technology is essential for a modern infrastructure of state of the art research centers. This is not just related to technologies for computation / storage and for the application of VR techniques for simulation and training but also for the IT infrastructure that enables distributed research groups to collaborate efficiently.

In this paper some novel solutions from different IT areas at the Forschungszentrum Dresden-Rossendorf are shown which will be discussed in more detail during this workshop. One solution addresses the use of GPUs for parallel computing, two examples show the application of VR technology for the simulation of a large beamline and the training of firefighters and another section shows the concepts and implementation aspects of an infrastructure for virtual institutes.

References

1. H. Schulz, J. Kelling, K.-H. Heinig, B. Liedke, G. Ódor, N. Schmeisser "Computing the KPZ equation Using GPU Acceleration". *Proc. of* the *International Workshop on Innovation Information Technology,* Vol. 1. Dresden, Germany, 2010
2. A. Arnold, N. v. Festenberg, S. Gumhold, P. Michel *"Komplexpraktikum Graphische Datenverarbeitung - Teilchenbeschleuniger"*. TU Dresden, Faculty of Computer Science, WS 0910
3. N. Schmeisser "A Virtual Firefighting Simulator". *Proc. of the International Workshop on Innovation Information Technology,* Vol. 1. Dresden, Germany, 2010
4. U. Konrad, N. Schmeisser "Identity Management in Large Research Facilities", *Proc. of the German-Russian Workshop on Innovation Information Technology,* Vol. 1. Ufa, Russia, 2009

Joy on Error - A critical view on technical artefacts

Volkmar Richter
Anhalt University of Applied Sciences.
Köthen, Germany
e-mail: v.richter@inf.hs-anhalt.de

Abstract

Miniaturisation and worldwide communication opens a wide area of prospective applications. Problems to develop secure software applications and examples of difficulties with the combination of hardware and software leads to the question: "how to ensure safety, usability as well as economic and ergonomic solutions in the future".

1. Introduction

Artefacts are the target and means of human work. Technological artefacts get a new level through combination of hardware and software. Substitution of mechanical functions by software may cause unprecedented failure rate. This paper identifies the main hazards and shows options for avoidance.

2. Technical artefacts

2.1 Actual development

The term was chosen intentionally. We are dealing with the current changing of technology. The hardware and software are converging and much of the former hardware functionality has been moved to the software level. This shift causes problems in the field of safety aspects. Of course it has advantages for producers and results in higher comfort for the user, etc.

The close coupling of hardware and software gets a new quality. Virtual objects replace real objects.

2.2 Software

That means programs, data, in short, everything that is able to influence the control of a computer system as an finite state machine. Let us look later at the difficulties in developing accurate software.
Using Software has a couple of problems.

- Digital computer technology is limited to specific number formats.
- Numbers are limited in precision, division by zero, overflow or rounding problems are the consequences.
- Processing time is not to disregard.

2.3 Hardware

Hardware also has special properties. Today's industrial mass production ensures low price and high quality. Lossless copy, like software is impossible, so that tolerances and testing steps like follows must be provided:

- Components Test
- Integrations Test
- System Test
- Roll Out Test
- Performance Test
- Load Test
- Crash Test

2.4 Interface crisis

Hardware with embedded software determines further problems by:

- multi-functional digital media
- historical keyboard genesis
- console's development
- desktop computing
- bottleneck Human-Computer-Interface
- natural operating interfaces
- invisible proactive computers
- increasing role of virtual objects

Therefore usability and ergonomic design gets a central role in research.

3. Errors

3.1 Error, fault and failure

Human errors, mostly caused by errors in the preparation, are resident as fault from beginning in the system, but they results only in a failure by combination of circumstances. Sometimes the compensation of two or more faults will be found (masquerade).

3.2 Computer designer and user

In ancient time of electronic computing were designers, builders and operators of the computer often identical. Errors were usually discovered quickly and corrected due to the small program size and familiarity of operators to technology.

Today's global network is based on a technocratic understanding of computer science. The risks arising from globalization are often eliminated.

3.3 Test and Software aging

Software as a lot of bits and bytes does not change by time. But social and technical environment will develop (relative aging). Otherwise programs maintenance will extend the number of errors (absolute aging). [1]

Edsger W. Dijkstra says in 1969: “Program testing can be used to show the presence of bugs, but never to show their absence!” [2]

Totally testing is unable; hence tests should start as early as possible. And bugs are not uniformly distributed, there are clusters of bugs. Testing depends also from environment. (Atomic power plant vs. website). Systems without failures don’t meet mandatory user’s imagination. [3]

Successful testing decreases costs of development and maintenance. But it is difficult to estimate probability of failures and the resulting costs, damages or injuries.

4 Global networking

4.1 Internet

The Internet has made networking many things that worked previously autonomous. But using a sprat to catch a mackerel we ignore or accept the risks like viruses, malware, spying by enemies or (legal) by government.

4.2 Social Networks

The internetworking enables global personal communications. New forms of learning and communication were developed, Google, Wikipedia, Twitter, etc. In the same manner we are able, regardless of distances, to communicate with persons or technical artefacts (avatars), learn about their condition and to intervene in the control of them.

4.3 Internet of Things

Various activities make it possible to equip things with IP addresses as a precondition to participate in global communication. The RFID technology was triggered by hype because it suddenly seemed very easy and cheap to achieve in the near-field interconnection and the automatic identification. [4]
Thus objects get new qualities like:

- Communication and co-operation
- Identification and addressability
- Sensor technology
- Actuator technology
- embedded information processing
- Localisation
- User interface

Partial successes, such as remote control of network components powered this idea. But we should not forget that network components already include the necessary communication hardware.

4.4 Ubiquitous Computing

The idea that computer science components could be almost anywhere has fascinating aspects. Total costs of ownership we like to forget like political and psychological aspects. Only the power supply seems to be a problem to solve.

The expectations reach form intelligent traffic control, efficient energy consumption to ambient assisted living. New services are planned for insurance companies, health care, warranty and costumer care. We are promised by "augmented reality" to get additional information by mesh up different internet sources.

Future challenges are:

- scalability
- arrive and operate
- interoperability
- energy budget
- communication bandwidth
- discovery
- software complexity
- narrow resources
- extensive software infrastructure in the web
- volume of data
- intelligent interpretation of data
- security and privacy protection:
- fault tolerance
- high level API on a low power device

But this technology has political, design and social dimensions. Personal information may be generated automatically and without consent and owners knowledge distributed. This may be potentially harmful and restricting individual freedom.

4.5 Dependency and Sovereignty

Data, automatically collected and interpreted by machine, may have a significant economic or social value. Who is the owner of this data and what ethical and legal framework should be valid. Dependency form technology (electricity, internet) is another risk and things which are controlled from a distance, may lead us in undesirable dependence and we will lose our self-determination. Our own smart objects were not always working as we wish, but like them (or their programmers) believe that it would be best for us.

Figure 1: Control rooms changes from panels to screen

5. Design Strategies

If you document it, it's not a bug -- it's a feature. Anytime a game in development crashed -- no matter how badly or bizarrely -- witnesses would invariably turn to the frustrated programmer, shrug, and calmly say "document it." [5]
Seriously, documentation is hard to do but best investment in the future.

The automation of various functions that were previously user-transparent, leads to an uncertainty for users, since the underlying algorithms

- are non-transparent,
- contain design errors
- are implemented incorrectly.

Software engineers have been concerned for years with this subject, with ergonomics and usability and in recent times additional with joy of use (the emotional satisfaction). [6]
For safety-critical applications exists techniques for secure operation, redundancy, error detection and recovery. The increasing number of networked heterogeneous systems and even for seemingly non-critical applications, need also methods to give them resilience. Otherwise exceed the cost of maintenance and troubleshooting each affordable level.

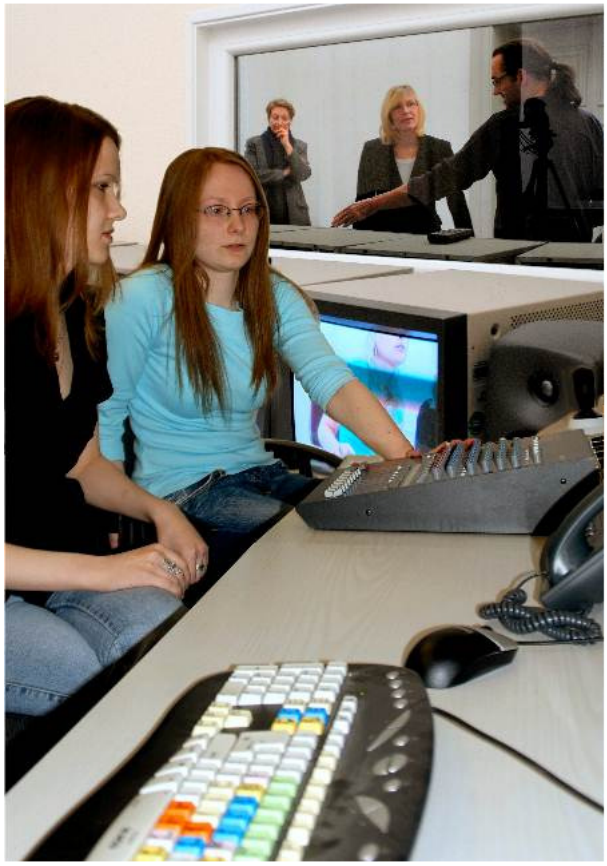


Figure 2: Human Centred Computing Laboratory

6. Conclusion

Resilience of software, the ability to continue tasks even after failure events with reasonable results will become more and more meaningful in a networking universe. Out department uses actual an Human Centred Computing Laboratory and is developing mobile instrumentation for it, to include also fixed installations found in power plants or in facility management.

Acknowledgements
This paper and related slides will be continued at www.inf.hs-anhalt.de/~volkmar/Joy_on_Error

References

1. Parnas D. L. “Software Aging”. *http://imps.mcmaster.ca/courses/SE-2A04-99/agingclass.slides.pdf (found 20/06/2010 09:35)*
2. Ercoli P. (ed.) “Software Engineering Techniques” *Rome, Italy, 27th To 31st October 1969, http://homepages.cs.ncl.ac.uk/brian.randell/NATO/nato1969.PDF (found 20/06/2010 09:45)*
3. Spillner A.; Linz T. „Basiswissen Softwaretest: Aus- und Weiterbildung zum Certified Tester - Foundation Level nach ISTQB-Standard“ *dpunkt-Verlag, Heidelberg, 2006.*
4. Mattern F.; Flörkemeier Ch. „Vom Internet der Computer zum Internet der Dinge“ *Informatik-Spektrum (9 February 2010)*
5. accredited to Mattel (Toy Brand) *http://intellivisionlives.com/bluesk...pace.html#hawk (found 20/06/2010 14:19)*
6. Palatini K.; “Joy of Use…”*3rd International Workshop “Innovation in Information Technologies”, Dresden, 2010*

On Adaptive Integration of Web Data Sources into Applications*

Roman Khazankin
Distributed Systems Group
Information Systems Institute
Vienna University of Technology
Vienna, Austria
e-mail: khazankin@infosys.tuwien.ac.at

Schahram Dustdar
Distributed Systems Group
Information Systems Institute
Vienna University of Technology
Vienna, Austria
e-mail: dustdar@infosys.tuwien.ac.at

Abstract[1]

Software applications often use data from web data sources, employing different integration techniques. The choice of a technique is made by an analyst or a developer traditionally. As non-functional characteristics of a web source and the business requirements of the application can change over time, the technique, that fits the situation best, might also vary. The manual control of the large number of integrated sources can be cumbersome as their characteristics might change independently. This paper gives an initial idea of (i) how the choice of a technique can be automated and (ii) how to make the integration adaptive to the environment.

1. Introduction

Software applications often use data from web data sources. This data can be used by applications employing different integration techniques. The choice of a technique impacts non-functional properties of utilized sources as from the perspective of the ultimate application. Because of the web's nature, the non-functional characteristics of web sources can change over time, so the choice of the most appropriate technique for data integration might differ. This is shown in motivating scenarios later in the paper.
The selection of data integration technique is driven by the business requirements on the one hand, and by the characteristics of the integrated data sources on the other hand. Such a choice is made by an analyst or a developer traditionally. However, with proliferation of online data sources, applications tend to use more remote sources [5]. If the application integrates a large number of independent web sources, the manual control of integration techniques can be cumbersome.

This paper shows how (i) the choice of technique can be automated and (ii) how the integration can be made adaptive to the environment therefore allowing for seamless switching of the technique to the most appropriate one according to specified criteria.
The approach's scope is applications that continuously use data from web sources. It does not address any scenarios where integrated data can be changed, so these changes need to be propagated back to the origin sources.

The paper is organized as follows: Section 2 describes the main integration techniques, Section 3 explains the factors that influence the choice of an integration technique, Section 4 depicts the motivating scenarios for the proposed approach, Section 5 describes the approach formally, Section 6 describes the basic architecture for the approach, Section 7 discusses the related work, Section 8 concludes the paper.

2. The techniques

The three main techniques used for integrating data are data *consolidation*, data *federation*, and data *propagation* (depicted in Fig. 1). All techniques are aimed to provide data from one or multiple origin sources in a specific form to the data consumer.
Data *consolidation* captures data from multiple sources and stores it into a single persistent target data store.
Data *federation* provides a single virtual view of one or more data sources. When the query is posed against federated view, it is forwarded to the origin sources. The responses are then combined and returned as a result.

---

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

*** This work was supported by the Vienna Science and Technology Fund (WWTF), project ICT08-032.**

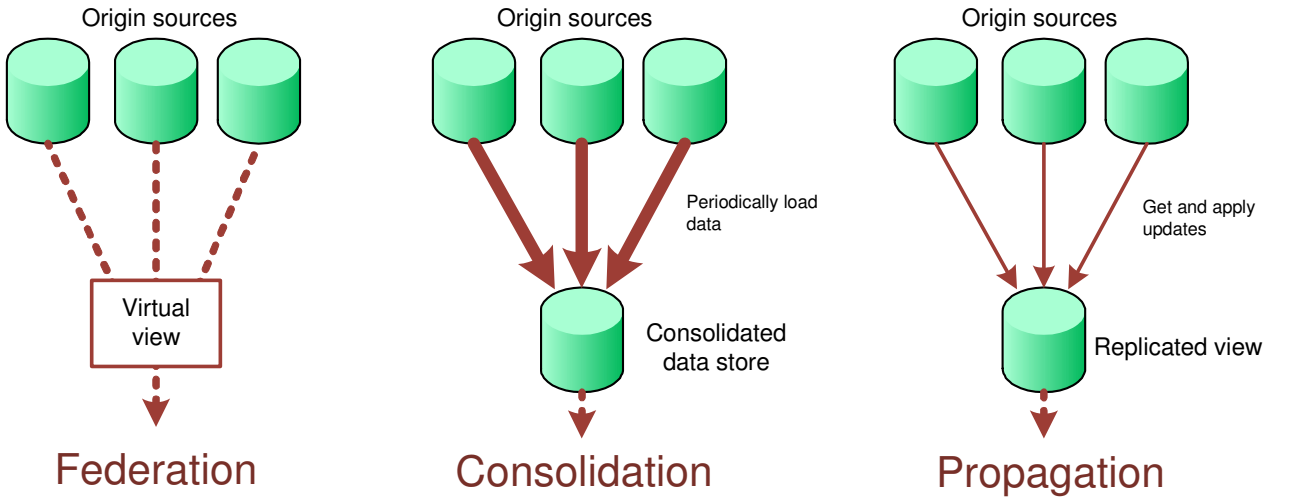


Figure 1: Main data integration techniques

Data *propagation* captures the changes in the source system and applies them to the replicated data store.
Federation provides real-time data and requires no additional storage, but makes the data consumer dependent on the origin sources' Quality of Service (QoS) properties and increases workload on them proportionally to the request frequency and requested data volume.

Consolidation guarantees QoS, but provides data with higher timeliness, requires additional storage and periodical workload on both origin and consumer systems proportionally to the requested data volume and the refresh rate.

Propagation guarantees QoS and provides near real-time data, but requires additional storage and increases workload proportionally to the update ratio. Propagation also requires tighter coupling as the access to the change log is required.

As a simple example, consider that company *First, Inc.* uses products from its partner, *Second GmbH*, to provide some services to their customers. To confirm its ability to provide the service to a customer, *First* needs the confirmation of the appropriate amount of products available from *Second*. Let *Second* have the data source with this information exposed. *First* needs to integrate this source into its system to use this information. As shown in Fig. 2, the choice of integration technique impacts the source's characteristics from the perspective of the ultimate application.

3. The technique choice

The choice of most suitable technique is driven by such concerns as timeliness of data, data volume, network latency, throughput, and availability of DS on the one hand, and by the corresponding requirements of consuming application (like latency or throughput) on the other hand. The examples of concerns and requirements regarded are listed in Table 1.

Table 1. Examples of data source offerings and data integration requirements

Offerings	Requirements
Timeliness	Timeliness
Latency	Latency
Throughput	Throughput (Request ratio)
Availability	Availability
Change ratio	Offline usage capability
Size of the source	
Update log availability	**General criterion**
Data storing permission	Disk space saving
	Workload saving
	Data transfer reduction

Different goals imply different requirements. For example, if the integrated source is frequently used by a few users at a time, e.g., for decision making, then the latency will be crucial; if the source is used to provide some information on company's website, then the throughput might be more important.

Enabling the analyst to set the priorities, we can automate the selection of most suitable technique. Having the priorities and/or restrictions defined, we can calculate the actual value for each criterion for each technique, and, finally, choose the technique which has the most satisfying calculated values according to priorities and/or restrictions.

4. Motivating scenarios

In this section we show the scenarios of applying the adaptive integration with explicit benefits.

4.1. Change of offerings

Extending the scenario from Sec. 2, let *First* receive some of the orders by phone, so it desires to confirm the availability in a short period of time (e.g., 3 seconds) to increase customer satisfaction. The required timeliness of data is 1 hour at maximum. Let *Second* have 10000 products which First uses. Usually *Second* provides data with mean latency=1 sec, so *First* employs federation which satisfies the requirement.
Now consider that network problems arise for two days and from the *First*'s perspective the latency increases to 5-10 sec, which hampers its business. Using the adaptive technique, the system can automatically detect that, for example, every-hour consolidation becomes more appropriate for the current situation, and can switch the technique, while the rest of the system would work as usual. After 2 days the network problems are fixed and adaptive technique switches back to federation to eliminate unnecessary periodical load.

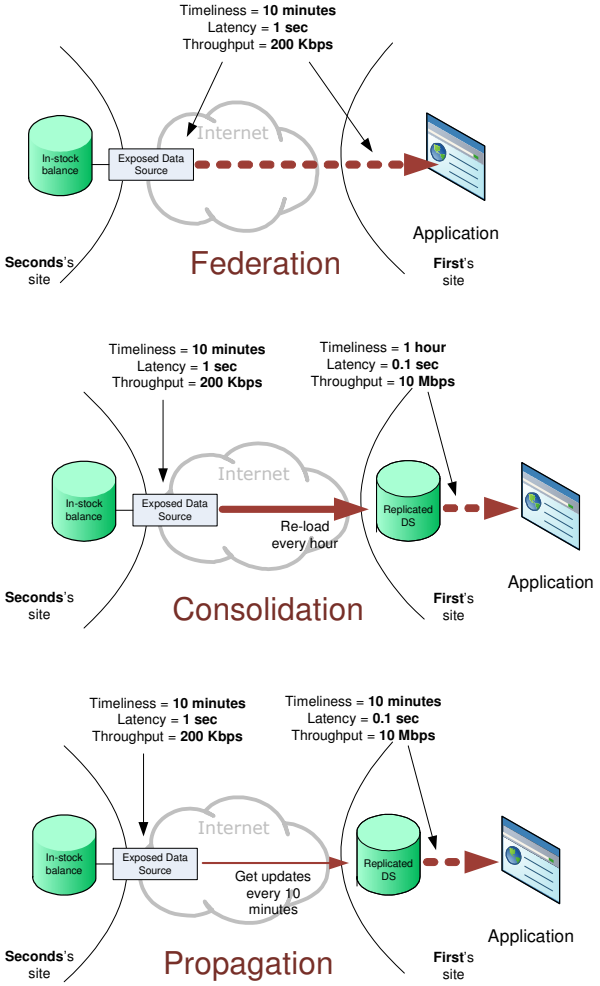


Figure 2. Impact of the technique's choice

While consolidation of one data source does not bring much overhead, it would make an impact when a large number of sources need to be integrated. Thus, in the case when an application integrates numerous external sources whose performance can change over time, the adaptation would optimize the disk space used by consolidated sources, the network load, and the workload caused by loads and transformations by automatically switching the technique used. Therewith, the manual control for a large number of integrated sources will be cumbersome.

4.2. Change of requirements

During the reporting period, the reports are generated much more frequently in a company. If these reports incorporate data from external sources (e.g., from the partners in virtual enterprise), then the throughput requirements become higher, so it might be more appropriate to use consolidation or propagation for integrating these sources, while normally federation is enough. The adaptive integration is able to seamlessly make decisions and automate such switching.

5. Formal specifications

In this section we formally outline the adaptive switching logic.
Let's assume that we regard N characteristics (or offerings) of a data source. Let $L(O) \in R$ - loss function that indicates how much penalty the characteristics $O = \{o_1,...,o_N\}$ result in being relevant for a short space of time Δt. L should be defined by the analyst and reflect the business requirements for the integrated source.
Let $C(T,P,O) = \{o'_1,...,o'_N\}$ define the influence of technique T applied with parameters P on the given set of offerings O. C should be defined by the system engineer and reflect the impact of the current technique implementation on data source characteristics as from the perspective of the ultimate application.
Let's assume that the offerings of the origin source are logged such as $o_i(l)$ - mean value of i_{th} offering in a l_{th} space of time, $i = \{1,..,N\}$. The length of each period is Δt. Let $R(l) = \{o_1(l),...,o_N(l)\}$.
Let $\Phi = \{l_1, l_2,....,l_M\}$ - a set of periods that form a basis for the decision making regarding the technique choice, such as last hour or last day.
Let $L_{fact}(l)$ - factual losses for l_{th} period, Ω - switching threshold.
Now, if there exist such $\{T,P\}$ that satisfy the following inequality, then the decision should be made in favour of technique T with parameters P:

$$\sum_{l \in \Phi} L_{fact}(l) - \min \sum_{l \in \Phi} L(C(T, P, R(l))) > \Omega$$

The rationale of this inequality is the following: it selects the technique which would have resulted in the least losses for the time period Φ, but the difference should be big enough to prevent from switching the technique too often, which is ineffective as the switching itself causes the loss. Ω reflects the cost of technique switch and depends on the particular environment and behavioral characteristics of the data sources and data integration tools used.

6. Adaptive integration architecture

This section describes the basic architecture of adaptive integration. As shown in the Fig. 3, the data from a web data source is provided to the end application via the ``proxy" source that acts as a wrapper using one of integration techniques underlying. The technique swapping is controlled by the adaptation module which makes decisions based on Loss function, Impact function, and the offerings log (see previous Section). The log is populated with the values either from web data source description via specification parser or from factual measurements via QoS measuring module.

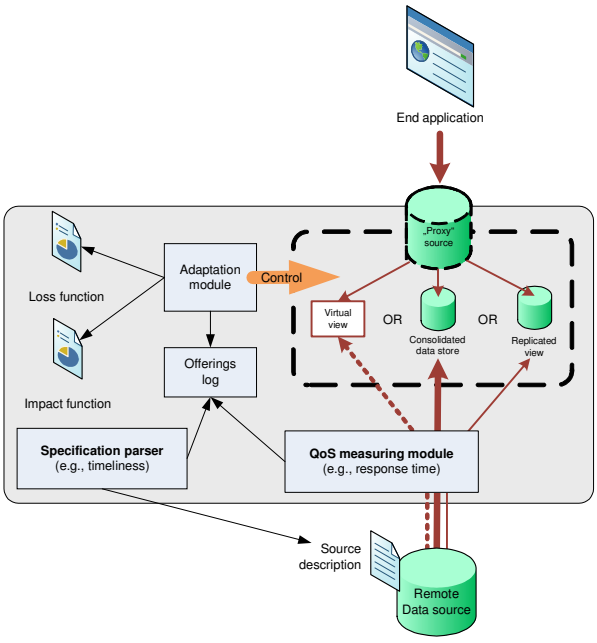


Figure 3. Adaptive integration architecture

The control module should check periodically if the currently used technique is still the best. When the decision about technique change is made, the tool implementing the new technique should be first prepared to serve the requests from the application, and only then the proxy can be re-bound to it. This ensures that the technique switch does not interrupt the application's work.

7. Related work

To the best of the authors' knowledge, no similar work has been done. The closest related area is adaptive databases or web pages caching [1,2,3,4].
The main difference is that the adaptation in those approaches aims to optimize the source system's performance, but not to fulfill the data consumer's individual business requirements.

8. Conclusion and future work

The initial ideas of adaptive data integration of web sources were presented in this paper. The problem was depicted and motivating scenarios were shown. Future work includes the development of real-world templates of impact and loss functions, the proper study of integration technique parameters, the analysis of applicable integration tools, prototype implementation.

References

1. C. Bornhoevd, M. Altinel, C. Mohan, H. Pirahesh, and B. Reinwald. Adaptive database caching with DBCache. *Data Engineering*, 27(2):11–18, 2004.
2. L.W.F. Chaves, E. Buchmann, F. Hueske, and K. Boehm. Towards materialized view selection for distributed databases. In *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology*, pages 1088–1099. ACM, 2009.
3. S. Guirguis, M.A. Sharaf, P.K. Chrysanthis, A. Labrinidis, and K. Pruhs. Adaptive scheduling of web transactions. In *Proceedings of the 2009 IEEE International Conference on Data Engineering*, pages 357–368, 2009.
4. A. Labrinidis, Q. Luo, J. Xu, and W. Xue. Caching and Materialization for Web Databases. *Databases*, 2(3):169–266, 2009.
5. Xiaofang Zhou, Shazia Sadiq, and Ke Deng. Data quality in web information systems. In *WISE '08: Proceedings of the 9th international conference on Web Information Systems Engineering*, Berlin, Heidelberg, 2008. Springer-Verlag.

Data Mining Technology on the Example of Bankruptcy Monitoring Problems

N.I.Yusupova
Ufa state aviation technical University I
Ufa, Russia
e-mail: yussupova@ugatu.ac.ru

T.V.Baronina
All-Russian State Distance-Learning Institute of Finance and Economics (ARDLIFE)
Ufa, Russia

G.R.Shakhmametova
Ufa state aviation technical University I
Ufa, Russia
e-mail: shakh@ugatu.ac.ru

E.O.Volik
Ufa state aviation technical University I
Ufa, Russia
e-mail: volik_e@mail.ru

Abstract[1]

The decision support system for monitoring bankruptcies, based on expert systems and data mining technologies, is discussed in article.

1. Introduction

An important aspect in solving bankruptcy problems is analysis and timely identification of the signs of fraudulent (fictitious and deliberate) bankruptcies as they cause the greatest damage in the sphere of company bankruptcy. The procedures for declaring a company a bankrupt and the process of identification of the signs of a possible fraudulent bankruptcy are defined by the law, however, this subject is an object of constant research as decisions are made under uncertainty and with inaccurate initial data. A special place in decision-making belongs to a human factor [1]. A big role in increasing the accuracy of managerial decisions is played by the monitoring of bankruptcies.

The article deals with a decision support system (DSS) developed by the authors of the article and intended for timely identification of the signs of fraudulent bankruptcies. The DSS makes use of the artificial intelligence technologies - those of expert systems and data mining.

Data mining represents a modern direction in the information systems area aimed at solving problems of decision support on the basis of quantitative and qualitative research of large arrays of heterogeneous retrospective data. One of the problems solved by means of data mining technology is forecasting [5,6]. The article looks at the possibility of using data mining to forecast finance indicators in the decision support system in monitoring bankruptcies and provides the analysis of specifics of finance indicators forecasting as well as some examples of prediction algorithms.

2. Decision support system for monitoring bankruptcies

The basis of the whole complex of techniques for the DSS are legally approved methodical instructions on accounting and analysis of companies' financial position and solvency so as to group companies depending on the level of risk of bankruptcy, as well as techniques for identification of the signs of fictitious and deliberate bankruptcy. These techniques are currently used by Russian auditors and arbitration managers [2]. To achieve these goals the authors propose the following general scheme of DSS in monitoring bankruptcies (Figure 1).

To develop the decision support system for monitoring bankruptcy the authors have used knowledge engineering, expert system technology and data mining technology.

The expert system technology [3,4] underlies two modules of decision support system in monitoring bankruptcies:

1. the module for grouping companies depending on the level of risk of bankruptcy;
2. the module for identification of the signs of illegal bankruptcy.

---

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

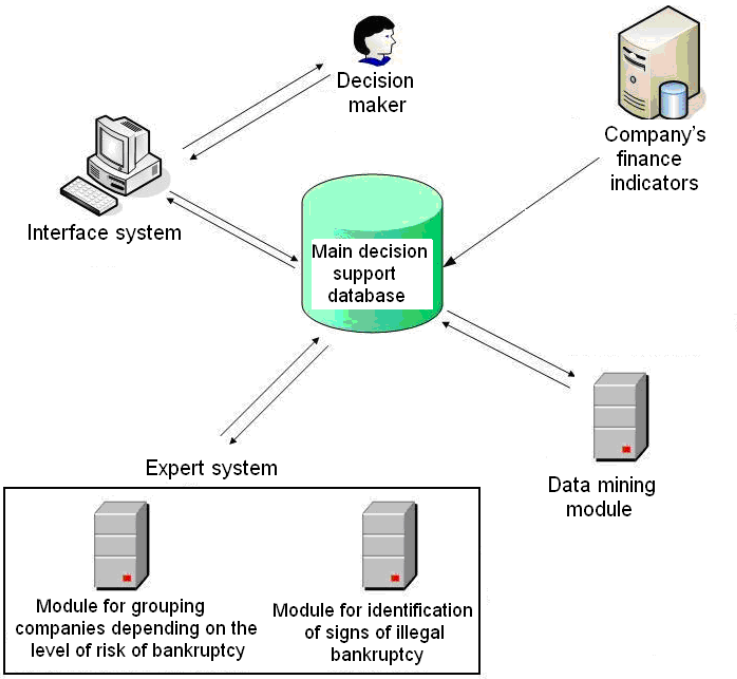


Figure 1. General scheme of DSS in monitoring bankruptcies

Another module of decision support system in monitoring bankruptcies is a data mining module. This module helps to solve problems which include cleaning the data for a qualitative forecast and predicting finance indicators of the company with the use of several prognostic model-building mechanisms, including self-teaching algorithms. The objective of this module is to identify negative trends in changing finance indicators as well as the possible signs of fraudulent bankruptcy based on the comparative analysis of the current and forecast in the data mining module finance indicators.

The primary, intermediate and resulting data are stored in the main decision support database. To keep the decision support system operating the primary data on the company is imported in the system either automatically or manually. Automatic data import can be done directly from the software 1C. Interaction between the DSS and user is made by means of an interface subsystem.

3. Data mining module in DSS

The problem which is solved by data mining module in DSS in monitoring a company's bankruptcy is a problem of forecasting finance indicators of the company. This problem can be seen as a problem of forecasting time series, as the data for the prediction of finance indicators are presented in the form of measurement sequences, collated in non-random moments of time. In contrast to the analysis of random sampling the analysis of time series is based on the assumption that successive values are observed in equal periods of time.

Like many other kinds of analysis the analysis of time series implies that the data contain a systematic component (generally including several components) and a random noise (error) which makes it hard to detect regular components.

An algorithm has been developed for predicting companies' finance indicators. It works as follows. Let's assume that as a result of transformation by the "sliding window" method we obtained a sequence of time counts**:**

$$X_{-n},\ldots\ldots,X_{-2},X_{-1},X$$

where X – current value. Forecast for X_{+1} is made on the basis of the built model. In order to forecast for value X_{+2} it is necessary to shift the whole of the sequence one count to the left so that the forecast X_{+1} made earlier could be included in the initial values. Then once again the algorithm for computing the predicted value will be started. X_{+2} will be calculated with regard for X_{+1} and so on according to the defined forecasting horizon. To debug the prediction algorithm it is necessary to define a forecasting horizon as well as the table's fields which must be input to make a forecast (to calculate the output field of the model).

Prediction of a company's finance indicators in DSS in monitoring bankruptcies can be done by means of a number of techniques such as a method of "sliding window" and neural networks.

In solving the problem of forecasting time series with the aid of a neural net it is required to input to the analyzer the values of several adjacent counts from the

initial set of data. This method of data sampling is called a “sliding window” (window - because only a certain area of data is highlighted, sliding - because this window “moves” across the whole data set). The efficiency of implementation considerably increases, if we do not sample the data out of a number of consecutive writings, but successively locate the data related to the specific position of the window in one writing. The values in one of the writing fields will be related to the current count and in other ones they will be shifted from the current count to the “future” or the “past”. Thus, transformation of the sliding window has two parameters: “depth of plunging” – the number of the “past” counts in the window and “forecasting horizon” - the number of “future” counts. It should be mentioned that for the boundary positions of the window (relative to the beginning and the end of the whole sampling) incomplete writings will be formed, i.e. writings containing empty values for the missing past and future counts. The transformation algorithm allows either to exclude such writings from the sampling (in that case for several boundary counts there will be no writings) or to include them (in the latter case writings will be made for all the counts available, but some of them will be incomplete).

To bring up a sliding window properly the data must be appropriately collated.

The neural network can be presented by the directed graph with weighted connections in which artificial neurons are the vertexes and synoptic connections – the arcs.

The multilayer neural net represents a set of neurons which make up layers. Within each layer the neurons are in no way linked with each other, but they are linked to the neurons of the previous and subsequent layers. The information is transferred from the first to the second layer, from the second - to the third, etc.

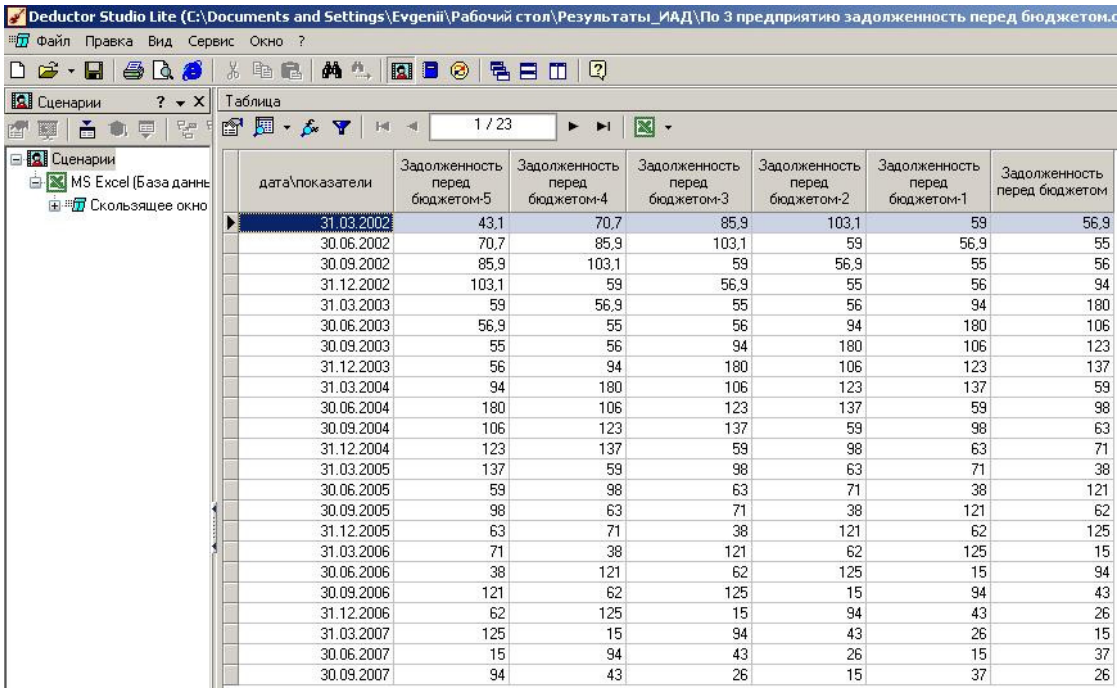
With the use of the neural net the forecasting problem can be set in the following way: to find the best approach of the function defined by the final set of input values (teaching examples). In our case the neural networks help to solve the problem of recovery of the missing values as well as prediction of finance indicators of the company being analyzed.

4. Software implementation of the data mining module

The software implementation of the data mining system to forecast finance indicators of the company is by means of the analytical platform Deductor 5.0.

Analysis of the primary data is made step-by-step. Cleaning, transformation and forecasting of the data is done individually with each time series of the company’s finance indicator. Let’s look at the whole of the data mining process on the example of the company’s budget indebtedness.

In forecasting time series by means of the neural net it is required to input the values of several adjacent counts from the source data set. This method of data sampling is called a sliding window**.** The efficiency of implementation considerably increases, if we do not sample the data every time from a number of consecutive writings, but successively locate the data related to the specific position of the window in one writing. Figure 2 shows the application of this method. Experiments show that for “sliding window” the optimal value of the depth of plunging is 5.



дата\показатели	Задолженность перед бюджетом-5	Задолженность перед бюджетом-4	Задолженность перед бюджетом-3	Задолженность перед бюджетом-2	Задолженность перед бюджетом-1	Задолженность перед бюджетом
31.03.2002	43,1	70,7	85,9	103,1	59	56,9
30.06.2002	70,7	85,9	103,1	59	56,9	55
30.09.2002	85,9	103,1	59	56,9	55	56
31.12.2002	103,1	59	56,9	55	56	94
31.03.2003	59	56,9	55	56	94	180
30.06.2003	56,9	55	56	94	180	106
30.09.2003	55	56	94	180	106	123
31.12.2003	56	94	180	106	123	137
31.03.2004	94	180	106	123	137	59
30.06.2004	180	106	123	137	59	98
30.09.2004	106	123	137	59	98	63
31.12.2004	123	137	59	98	63	71
31.03.2005	137	59	98	63	71	38
30.06.2005	59	98	63	71	38	121
30.09.2005	98	63	71	38	121	62
31.12.2005	63	71	38	121	62	125
31.03.2006	71	38	121	62	125	15
30.06.2006	38	121	62	125	15	94
30.09.2006	121	62	125	15	94	43
31.12.2006	62	125	15	94	43	26
31.03.2007	125	15	94	43	26	15
30.06.2007	15	94	43	26	15	37
30.09.2007	94	43	26	15	37	26

Figure 2. Program window of implementation of sliding window method

The neural net structure for predicting the company's budget indebtedness has the form 5-2-1. The parameters and graph of the neural network are shown in Fig.3

.

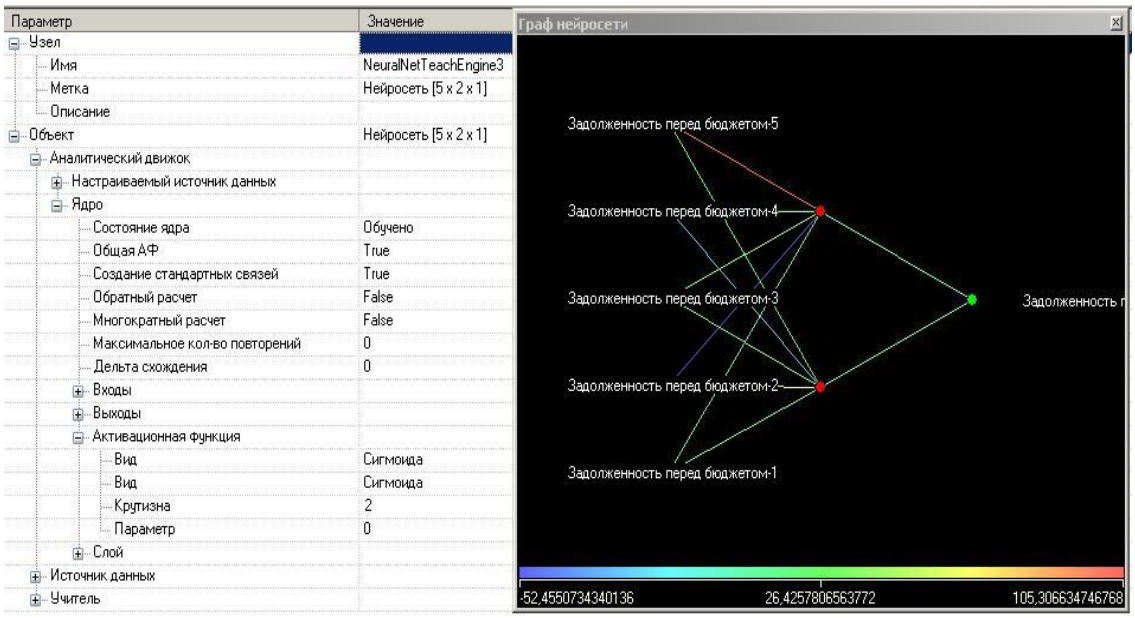
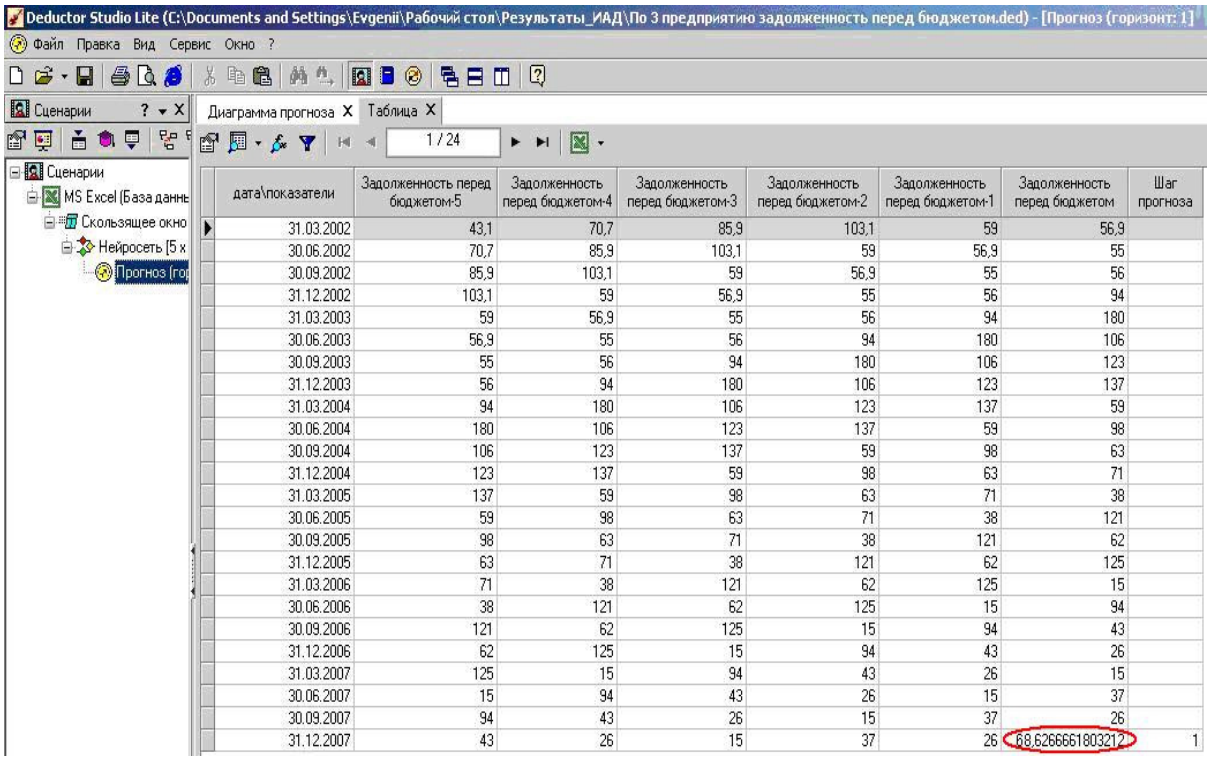


Figure 3. Program window of parameters and graph of neural net

The final stage in the program implementation of data mining on the example of the company's budget indebtedness is forecasting. Forecasting in the analytical platform Deductor is implemented by means of the data processing master capable of forecasting time series. The program window with the prediction results is shown in Figure 4.



дата\показатели	Задолженность перед бюджетом-5	Задолженность перед бюджетом-4	Задолженность перед бюджетом-3	Задолженность перед бюджетом-2	Задолженность перед бюджетом-1	Задолженность перед бюджетом	Шаг прогноза
31.03.2002	43,1	70,7	85,9	103,1	59	56,9	
30.06.2002	70,7	85,9	103,1	59	56,9	55	
30.09.2002	85,9	103,1	59	56,9	55	56	
31.12.2002	103,1	59	56,9	55	56	94	
31.03.2003	59	56,9	55	56	94	180	
30.06.2003	56,9	55	56	94	180	106	
30.09.2003	55	56	94	180	106	123	
31.12.2003	56	94	180	106	123	137	
31.03.2004	94	180	106	123	137	59	
30.06.2004	180	106	123	137	59	98	
30.09.2004	106	123	137	59	98	63	
31.12.2004	123	137	59	98	63	71	
31.03.2005	137	59	98	63	71	38	
30.06.2005	59	98	63	71	38	121	
30.09.2005	98	63	71	38	121	62	
31.12.2005	63	71	38	121	62	125	
31.03.2006	71	38	121	62	125	15	
30.06.2006	38	121	62	125	15	94	
30.09.2006	121	62	125	15	94	43	
31.12.2006	62	125	15	94	43	26	
31.03.2007	125	15	94	43	26	15	
30.06.2007	15	94	43	26	15	37	
30.09.2007	94	43	26	15	37	26	
31.12.2007	43	26	15	37	26	68,6266661803212	1

Figure4. Program window with the prediction results of the company's budget indebtedness in the following reporting period

Each of the finance indicators has its own prediction algorithm that includes the size of the step of the sliding window, neural net structure, the form of the activation function and its steepness (Table 1).

Table 1. Algorithms of data mining application to forecast company's various finance indicators

Company's finance indicators	Depth of plunging of sliding window	Form of the activation function	Value of activation function	Neural net structure
Fictitious assets, i.e. patents, licenses, trade marks	3	Sigma form	0,7	5-2-1
Fixed assets	5	Sigma form	1,9	5-4-3-2-1
Long-term financial investments	3	Sigma form	1	3-2-1
Total of non-working assets	5	Sigma form	1,03	5-2-1
…	…	…	…	…
Reserves of forthcoming expenses and payments	5	Sigma form	1	5-2-1
Total of short-term liabilities	5	Sigma form	1	5-2-1
Liabilities balance	5	Sigma form	0,9	5-2-1

Examples of the structures of used multilayer neural nets are shown in Figure 5.

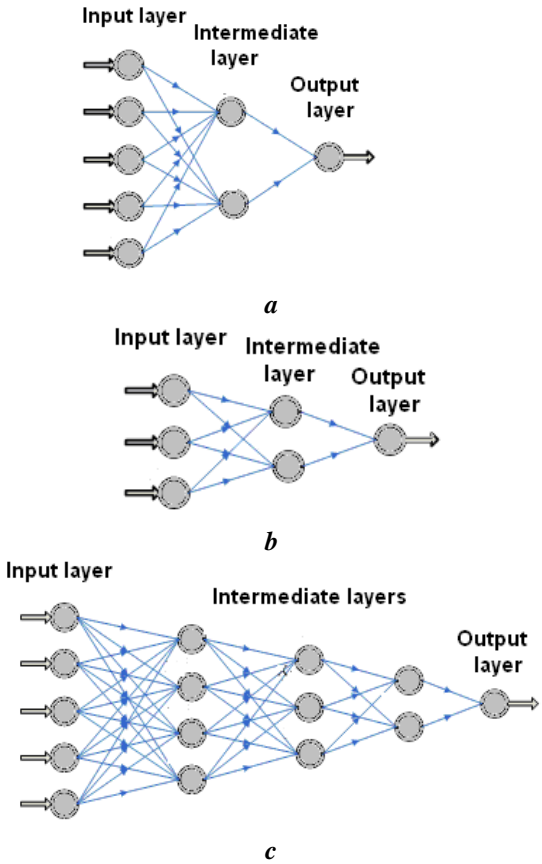


Figure 5. Examples of used multilayer neural nets:
***a* – graph of neural net 5-2-1;**
***b* - graph of neural net 3-2-1;**
***c* - graph of neural net 5-4-3-2-1.**

5. Conclusion

The article examines a method for implementation of the data mining module in the decision support system in bankruptcy monitoring. Analysis has been made of the requirements for the data of the decision support system and models have been selected for the data mining implementation. The article also provides a description of the way the data mining module for predicting finance indicators of the company is implemented on the basis of the analytical platform DEDUCTOR by means of the sliding window method and multilayer neural nets. For each of the predicted finance indicators there has been developed a separate prediction algorithm.

Acknowledgement

This research has been partially supported by grants 09-07-00408-a (Russian Foundation for Basic Research), 09-02-84-201 (Russian Humanitarian Scientific Fund), research work "Research of intellectual technologies for decision-making support and control of complex social and economic objects" (Ministry of Education of the Russian Federation) and by grant of President of Russian Federation № 65497.2010.9 for Leading Scientific Schools.

References

1. Timofeev O.V. “Bankruptcy procedures as an instrument of take-over raid attacks”. In: Trans-regional forum "Actual questions of economic security. Take-over raids as a barrier to Russia’s industrial growth". Ufa, 2007.
2. Yusupova N.I., Volik B.O. "Bankruptcy monitoring with the use of intellectual data analysis". Vestnik USATU, 2008; Vol. 10, 2(27): 71-79.
3. Jackson P. "Introduction to expert systems". Williams, Moscow, 2001.
4. Popov E.V. "Statistic and dynamic expert systems". Finance and statistics, Moscow, 1996.
5. Barsegyan A.A., Kupriyanov M.S., Stepanenko V.V., Kholod I.I. "Methods and models of data analysis. OLAP and Data Mining". BHV-Petersburg, Saint-Petersburg, 2004.
6. Duk V., Samoilenko A. "Data Mining". Peter, Saint-Petersburg, 2001.

Semantic Modeling. Computational Models of the Concepts

Viacheslav Wolfengagen
Dept. of ACS & IT
Institute Jurinfor-MGU/NRNU MEPhI, Moscow Russia
e-mail: vew@jmsuice.msk.ru

Abstract[1]

As known, computing allows maintenance of the constructions corresponding to the needs in comprehended models which enable the computational experiment. Computing takes the enabling of such a model which corresponds to a `theory'. The parameters of the model are systematically changing and correspond to `entries' for the inputs. The parameterized family of its `behaviors' is used as the `outputs' of the model. These -- being computed, -- behaviors are matched with the real behavior of the problem domain object under the real inputs – prototypes of model inputs. In case of conclusion that the model satisfactory reflects the features of the real object-prototype, then the real experiment can be replaced by the computational one, which is executed using the model-image.
Keywords: semantic modeling, applicative computing, variable domains

1. Introduction

The current scientific revolution first of all caused a significant shift of the method of science. A former way of research -- from setting up a hypothesis to development of equipment and experimentation, and then to experimental data analysis, -- is replaced by the new way. This new way looks differently -- from formulating a hypothesis to the data base search for an answer. This clearly underlines the primary place of databases in the infrastructure of a science. A database can contain the raw data, the results of computational analysis or modeling, and the result of some special data structuring as well.

A modern trend in science uses database publication. But the structure of scientific database significantly differs from, say, business database. For instance, the data in store can be just a program code which is used to optimize the queries. In addition, this is not always enough to involve in a formulation of a distributed query just the operations of relational algebra. The operations with the constructions of programming language used in a science can be added with some other operations as well. The stored data are more heterogeneous, e.g. a processing of data with the documents' fragments can be involved. The metadata can be used as an information of the stored data, their features etc. Metadata give not only just the suitability for user, but a means for data integration and usability of Web-services which have an ability to construe the derived data. In general, modern science needs the interfaces with very large and complicated data sets. There is a need to manipulate, visualize and interpret such sets of data.

The more important for a separately given science is an integration of the theory, experiments and models. As known, computing allows maintenance of the constructions corresponding to the needs in comprehended models which enable the computational experiment. Computing takes the enabling of such a model which corresponds to the `theory'. The parameters of the model are systematically changing and correspond to `entries' for the inputs. The parameterized family of its `behaviors' is used as the `outputs' of the model. These -- being computed, -- behaviors are matched with the real behavior of the problem domain object under the real inputs – prototypes of model inputs. In case of conclusion that the model satisfactory reflects the features of the real object-prototype, then the real experiment can be replaced by the computational one, which is executed using the model-image [1].

2. Basic fundamentality of computation

When involving in an area of computing pay attention, what are the questions -- from a supply of generic ones, -- which have no determined answer:
-- what is a `data'?
-- what is a `computation'?
This indeterminacy can be perceived by different ways, for instance in the following context:
(1) either both of the concepts co-exist, i.e. they are complementary, or

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

(2) one of these concepts is excessive.
In modern time, in connection with obtaining the answer for question (1), it may seem that both of the lines are working out relatively independent:
-- data bases, which are studying at a model level distinctly, but with a confident intersection that they operate with the `data sets',
-- algorithms and programs, which give rise to the `model of computation(s)'.
Nevertheless the question (2) can be more attractive and having at least one positive answer which is given by the *combinatory logic*.

It contains the special `constants' referred as *combinators*, each of them is assumed as a `building block' taking part in a `construct of computation'. Hence combinator represents the least sense construction of a `computation'. Therefore the combinators are `given', i.e. they are given data. On the other hand every combinator is a `computation' the same time, i.e. it is a process. This means that combinators have the surprising duality having being both the data and computations: data-computations. No one of these aspects is allowed to be ignored.

2.1. Explanatory systems of computations

Is there any explanatory system which gives an outlook of combinators' behavior, these intriguing data-computations? This system does exist and it is known by the name of `theory of computation'. A mathematical mapping of the outlook of computations with combinators is that they could be acknowledged as the constants-processes, and both the constants and processes are considered in a mathematical sense of these terms.

2.2. Theory of computation

In a computation theory the behavior both of combinators and of the various constructions which can be combined from them are to be described. What is a scope of this theory? First of all this is a theory of meaning, i.e. each of its `objects' is assigned to the `meaning'. A set of all the assignments can compare the different possible `states' of computation. But we tend to the determined general outline of computations. This is not a reason to avoid an uncertainty but by a simple reason that we are not yet learned to get the profit from the undetermined computations.

2.3. Generic computations and their origin

Isn't it too early to decide that the computations are clearly known and their relatives are gathered in a couple which is ramified as `computing'? Is everything clear with an origin of computation? And has the computation an `origin' as it is considered as an artifact? Note that the recent years gave rise to the doubts in a purely artificial nature of computation and the information transfer is assumed attributive to the natural processes and structures.

2.4. Knowledge inheritance/acquisition

Usual schema of knowledge acquisition -- from older generation to younger, -- is violated as the information technologies are being involved. Both of the subjects -- an agent of knowledge transfer and its recipient, -- are within the "information field" which is rapidly evolving both in qualitative and quantitative respect. Conceptual framework and its structure are not yet stabilized to allow the step by step setting up of mastering in the fine details of a skill. The new concepts are generating which has to be learned, they induce a structure of dependencies demanding the rapid assuming. The known concepts can change their meaning and/or interpretation, their existing structural dependencies lose the previous meaning and inherit a new one.

The rate of changes is so high that both the knowledge agent and recipient are in a position of the learner which needs rapidly to learn a lesson given by a reality, and the advantage has not a person who has a deeper understanding of the matter, but such a person who neglects the former representation and accepts the new idea having a readiness namely to changes. The dynamic of changing the representations is becoming an object of analysis, and stable conceptual structures exist on the rights of invariants.

Constancy of concepts drops down its dominant position is becoming replaced by the variations of concepts. The expressing of dynamics of a concept needs not only the adequate language which gives a ground to do this, but the dissemination of these languages among the variety of consumers -- those who obtain the knowledge. This new language is yet in a stage of proliferation, but already has started to take that a universal role which was previously played by the mathematics in its classic meaning.

To express this idea some supply of the mathematical means have to be applied.

A *constant concept* can be determined by the postulate

$$(\eta) \quad \lambda x.Xx = X, \text{ where } x \notin X.$$

It would be better to write the expression for a "constant function"

$$(\lambda x.c)a = c$$

for $x \not\equiv c$ and any a. Writing

$$(\lambda y.yx)F = Fx,$$

we obtain the expression of a "variable function". Then by the rule of (ξ) we write down

$$\frac{(\lambda y.yx)F = Fx}{\lambda x.(\lambda y.yx)F = \lambda x.Fx,} \quad (\xi)$$

which gives the formulation on an extensionality (η) one more time. But not everybody takes this purely syntactic consideration as the satisfactory. Expressing an idea of `being a concept', i.e. of `being a notion' usually attracts the means of semantic modeling and assignments.

2.6. A theory of the constant concepts

Give an outline of a theory of constant concepts in accordance with the semantic considerations.

As for the concept C it is possible to write

$$\|\lambda x.Cx\| : \text{environment} \to (\text{individual} \to \{true, false\}),$$

then

$$\|\lambda x.Cx\|\rho d = \Lambda(\varepsilon \circ \langle \|C\|, \|x\| \rangle)\rho d =$$

$$= \varepsilon[\|C\|[\rho, d], \|x\|[\rho, d]]$$

$$= \varepsilon[C, d]$$
$$= C(d).$$

In a language of sets it means that

$$\|\lambda x.Cx\|\rho \subseteq C \in [C],$$

and for assignments $\rho \in I$ in the language of types it follows that

$$\|\lambda x.Cx\| : I \to C \to [\,] \text{ or } \|\lambda x.Cx\| : I \to [C].$$

Hence, it is derivable from above that

$$\|\lambda x.Cx\| = \Lambda\|Cx\| : I \to C \to [\,]$$

$$= \|Cx\| : (I \times C) \to [\,]$$

$$= \|Cx\| : [I \times C].$$

Now from

$$\lambda x.Cx = C,$$

it can be written that

$$\|\lambda x.Cx\| = \|C\| : I \to C \to [\,].$$

2.7. A theory of the variable concepts

To obtain a theory of variable concepts the results for the constant concepts above will be used with the difference that C is associated with a parameter denoted by T. This parameterization leads to the sets

$$H_T(I) = \{h \mid I \to T\}.$$

The newly obtained equalities can be re-written with T. In a language of the sets it means that

$$\|\lambda x.Cx\|\rho \subseteq T \in [T],$$

and for assignments $\rho \in I$ in the language of types it follows that

$$\|\lambda x.Cx\| : I \to T \to [\,] \text{ or } \|\lambda x.Cx\| : I \to [T].$$

It follows from above that

$$\|\lambda x.Cx\| = \Lambda\|Cx\| : I \to T \to [\,]$$

$$= \|Cx\| : (I \times T) \to [\,]$$

$$= \|Cx\| : [I \times T].$$

Now from $\lambda x.Cx = C$, it follows that

$$\|\lambda x.Cx\| = \|C\| : I \to T \to [\ ].$$

3. Semantic nets

The object of a semantic net is the structure with predicate and its arguments. The objects are packed into a *complete partial ordered* structure (c.p.o.) which, in terms of semantic nets is called as an *ISA-hierarchy*. Objects are involved in an interaction by means of the *operations* where they participate as *operands*. Execution of an operation leads to the «overpacking» of a c.p.o. Thus the interaction of objects is executed by means of the intermediators-operations. But in a couple the interaction with an *environment* is subordinated the laws of c.p.o. An interaction as itself is dropped down to the set of operations and the properties of this set should be studied. C.p.o. generates the content which is called as *intension*. The induced sets generate the configuration which is called as *extension*.

3.1. The notion of a problem domain dynamics

Those objects that reflect the dynamic of a problem domain have a rather complicated semantic structure, namely, the semantic net. A system of semantic representation both of the objects and of thei relations is as advanced that the object is simultaneously located in the two dimensions, in the two «realities» -- *actual* and imaginable, or, in other terms, *virtual* one. The imaginable world mainly and, for most of the nets, dominantly determines the behavior of the objects. But it is easy changing, and can be affected from outside so that the net will miss this action.

The objects representing a problem domain, take part in different linkages giving rise to semantic net. Only those objects that are specially selected out can exist and be taken into account in isolation, outside the net. The *individual* is an abstraction, an idealized representation of the isolated object. By the sense of this term, an individual is assumed as *undividable*. In practice, an idealized representation of the individual is not feasible, the objects are generated and exist only in interaction with each other, having been captured by the «sphere of influence» of other objects.

3.2. The problem of the "fast Smith"

Let us examine a simplest example which shows, how to represent the dynamic of a problem domain using the principles of comprehension. As it will follow, this process involves both the *concepts* and *individuals*. This example has the nature of a hint, but due to its generality, the fundamental effect of the provoking substitution, which has been illustrated, deserves a special mention as the task of the "fast Smith".

Example 1. Lets us represent the fragment of a mapped problem domain:

'The employee Smith is not married, but the employee Jones is married. Smith became married taking the name 'Jones' '.

Write down the frames and the definitions of the concepts:

def

EMPLOYEE.NAME[x] = [∀x. MARITAL STATUS

(arg : x : PERSON,
res :y : STATUS)],

def

NOT-MARRIED[x] = [∀x. MARITAL STATUS

(arg : x : EMPLOYEE.NAME,
res : 'not-married' : STATUS)],

def

MARRIED[x] = [∀x. MARITAL STATUS

(arg : x : EMPLOYEE.NAME,
res : 'married' : STATUS)],

def
F1 = MARITAL STATUS [
(arg : x : EMPLOYEE.NAME,
res :y : STATUS)],

def
F1e = MARITAL STATUS [
(arg : 'Smith' : NOT-MARRIED,
res : 'not-married' : STATUS)],

def
F2e = MARITAL STATUS [
(arg : 'Jones' : MARRIED,
res : 'married' : STATUS)],

NOT-MARRIED ISA EMPLOYEE.NAME ISA PERSON,
MARRIED ISA EMPLOYEE.NAME ISA PERSON,

F1e Instance-of F , F2e Instance-of F .

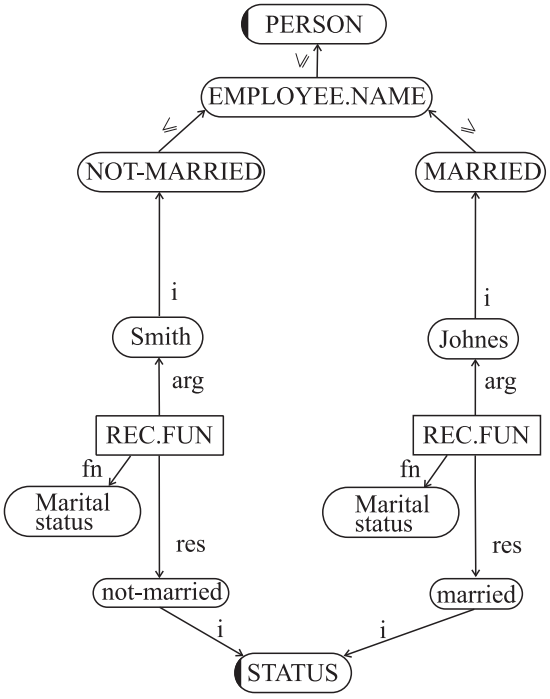


Figure 1 Recognition function MARITAL STATUS

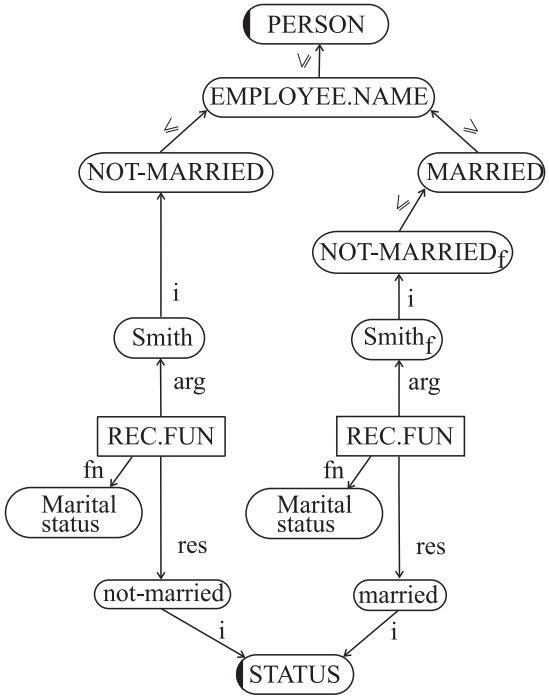


Figure 2 Employee Smith get married

This representation is reflected in Fig. 1. This figure gives the following qualitative framework of the behaviorr of individual 'Smith'. In its initial -- "old", -- state 'Smith' belongs to those employees that are not married. This means that the concept NOT-MARRIED, derived by applying the recognition function REC.FUN for status 'not-married' from the concept STATUS is maintained. In the same state 'Jones' belongs to those employees that are married. This means that the concept MARRIED, derived by applying the recognition function REC.FUN for status 'married' from the concept STATUS is maintained.

Let the events now are evolved along the mapping f generating the event

'Employee Smith is married (i.e. he get married)'.

The following frame-instantiation fixes this event:

def

F22e = MARITAL STATUS [

(arg : 'Smith$_f$ ' : NOT-MARRIED$_f$,

res : 'married' : STATUS)],

NOT-MARRIED$_f$ ISA MARRIED

This is represented in Fig. 2.

This figure gives the following qualitative framework of the behavior of individual 'Smith'. In its initial -- "old", -- state 'Smith' belongs to those employees that are not married. This means that the concept NOT-MARRIED, derived by applying the recognition function REC.FUN for status 'not-married' from the concept STATUS is maintained. Upon transition, having being in current -- "new", -- state 'Smith$_f$' belongs to those employees that were not married but now are married. This means that the concept NOT-MARRIED$_f$, derived by applying the recognition function REC.FUN for status 'married' from the concept STATUS is maintained.

Assume that upon get married Smith has changed his name by Jones:

'employee Smith get married taking the name 'Jones' '.

def

F23e = MARITAL STATUS [

(arg : 'Smith$_f$' : NOT-MARRIED$_f$,

res : 'married' : STATUS)],

AND EQ [

(arg1 : 'Smith$_f$' : NOT-MARRIED$_f$,

(arg2 : 'Jones' : MARRIED$_f$

res : 'true' : TRUE-FALSE)]

This is represented in Fig. 3.

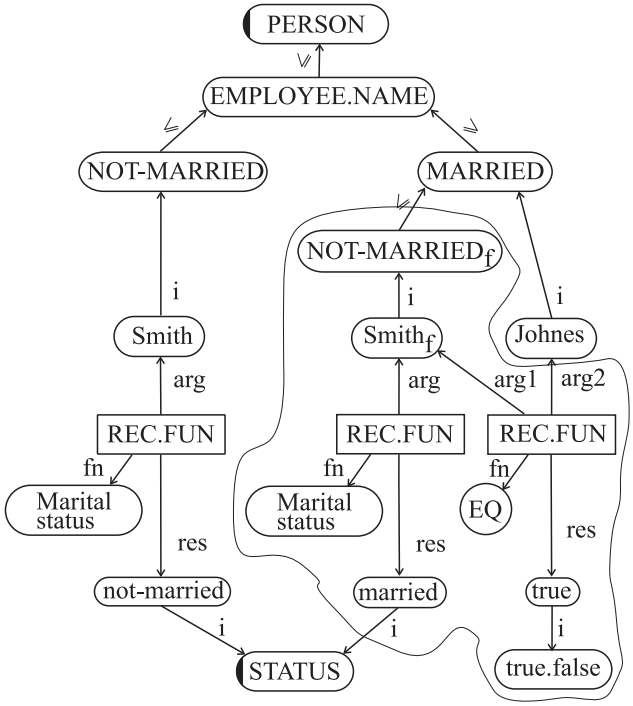


Figure 3 Employee Smith get married and changed the name by 'Jones'

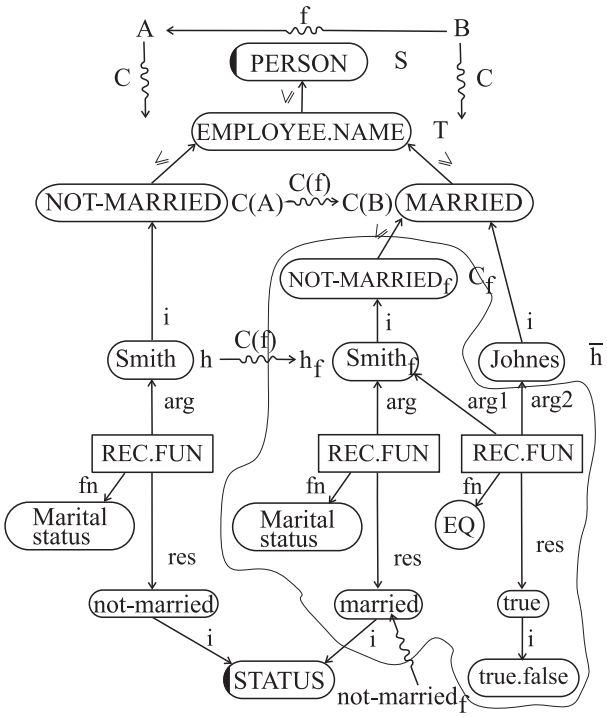


Figure 4 Dynamics of evolving the events, corresponding the situation 'employee Smith get married and changed the name by 'Jones' '

This figure gives the following qualitative framework of the behavior of individual 'Smith'. In its initial -- "old", -- state 'Smith' belongs to those employees that are not married. This means that the concept NOT-MARRIED, derived by applying the recognition function REC.FUN for status 'not-married' from the concept STATUS is maintained. Upon transition, having being in current -- "new", -- state 'Smith$_f$' belongs to those employees that were not married but now are married. This means that the concept NOT-MARRIED$_f$, derived by applying the recognition function REC.FUN for status 'married' from the concept STATUS is maintained.
Besides that, there is the individual 'Jones' belonging to those employees which are married.
This means that the concept MARRIED, derived by applying the recognition function REC.FUN for status 'married' from the concept STATUS is maintained. In the figure this function is just omitted.
At last, the situation where individuals 'Smith$_f$ ' and 'Jones' in a current state are "merged" is indicated by the logical predicate EQ. This predicate sets an equality of the individuals 'Smith$_f$ ' and 'Jones' in the new state which is under consideration.
Now Fig. 4 reflects a dynamic of evolving the events corresponding the situation
'employee Smith get married and changed the name by 'Jones' '.

This figure gives the following qualitative framework of the behavior of individual 'Smith'. In its initial -- "old", -- state 'Smith' belongs to those employees that are not married. This means that the concept NOT-MARRIED, derived by applying the recognition function REC.FUN for status 'not-married' from the concept STATUS is maintained. Upon transition, having being in current -- "new", -- state 'Smith$_f$' belongs to those employees that were not married but now are married. This means that the concept NOT-MARRIED$_f$, derived by applying the recognition function REC.FUN for status 'married' from the concept STATUS is maintained.
Besides that, there is the individual 'Jones' belonging to those employees which are married.
This means that the concept MARRIED, derived by applying the recognition function REC.FUN for status 'married' from the concept STATUS is maintained. In the figure this function is just omitted.
At last, the situation where individuals 'Smith$_f$' and 'Jones' in a current state are "merged" is indicated by the logical predicate EQ. This predicate sets an equality of the individuals 'Smith$_f$ ' and 'Jones' in the new state which is under consideration.
All of this have already been represented in the previous figure. A new feature is the edges with the waved line linking the "evolving of events" *f* from *A* to *B* with the frame. An area characterizing the cross-identification of the individual which changes its properties is bounded.
A dynamic reflected in this example will be written in more general form. The events in a problem domain

are evolving in the direction from A to B along the evolvent f:

$$f : B \rightarrow A$$

(pay attention to a reversed order!). A set of persons is denoted by S, and a set of employees names is denoted

by T, where $T \subseteq S$.

With this assumption the marital status is associated to a generalized concept $C \equiv H_T$ such that

$$H_T(A) \equiv C(A) \subseteq T, \qquad H_T(B) \equiv C(B) \subseteq T,$$

$$C(f) : C(A) \rightarrow C(B).$$

It can be noted that

$$C_f \subseteq C(B),$$

which is in a full harmony with an expected behavior of the model and intuition. The element wise consideration gives

$$h^- \in C(B), \qquad C(f): h \mapsto h_f .$$

Using the terms of a problem domain we conclude that

$f \equiv$ marriage with changing the name,

A is a problem domain 'before marriage', and B is a problem domain 'after marriage'. Besides that,

$S \equiv$ PERSON ,$T \equiv$ EMPLOYEE.NAME,

and

EMPLOYEE.NAME ISA PERSON.

In addition,

NOT-MARRIED ISA EMPLOYEE.NAME,

MARRIED ISA EMPLOYEE.NAME,

NOT-MARRIED_f : NOT-MARRIED → MARRIED.

An element wise consideration gives:

NOT-MARRIED_f : Smith ↦ Smith_f , Smith = Jones

etc. This may finalize a consideration of the Example 1.

4. Conclusion

The experiments have shown the applicability of both a theory of constant concepts and a theory of variable concepts to semantic modeling.

References

1. V. E. Wolfengagen. "Applicative computing. Its quarks, atoms and molecules". -- L.Yu. Ismailova (Ed.). -- Moscow: "Center JurInfoR", 2010. -- 62 p.

Applicative Models, Semantic Scalability and Specialized Calculations for Business Games in Jurisprudence

Larissa Ismailova
Institute for Contemporary Education "JurInfoR-MSU",
Moscow, Malaya Pirogovskaya st., 5, office 37, 119435, Russia
e-mail: larisa@jurinfor.ru

Sergey Kosikov
Institute for Contemporary Education "JurInfoR-MSU",
Moscow, Malaya Pirogovskaya st., 5, office 37, 119435, Russia
e-mail: kosikov.s.v@gmail.com

Abstract[1]

Herein suggested are the design solutions for the business games of the modeling type in jurisprudence as well as a design concept for specialized subject oriented mechanisms of inheritance their composition and the ways to integrate them in the applicative computing systems (ACS) and releases.
Keywords: applicative computing models, ACS, semantic equation, combinatory, abstract machine, functional languages, business games.

1. Introduction

Specificity of formulating semantic aspects of business games in jurisprudence manifests itself at every level of their designing [1]. Utilization of the means of semantic fine tuning is dictated by the essential openness of the business game description language and requirements of its further maintenance. Generally the subject matters of the juridical business games are difficult to formalize [2,3,4] and the modeling of the decision making situations must account for the conditions of: (1) truth-value gaps; (2) oversaturated and (3) volatile evaluations and mechanisms of evaluation.

The business games subject matters in jurisprudence are known to be described in the framework of differing terminological "systems". The latter oftentimes acquire various possible interpretations in the model and may change with time or under influence of other objects of the environment.

In the chosen subject matter majority of the terminological collisions spill beyond the doctrinal ones that has essential practical importance. It requires control over techniques of semantic valuation of the language phraseology. Thus emerges necessity to design specialized mechanisms of inheriting where the rule of inheriting itself turns out to be calculable.

Review of the contemporary sources indicates the lack of clear cut understanding of applicability of the traditional calculation algorithms resolving these problems. It makes conceptual modeling the main method of description of the difficult to formalize subject matters.

Development of the mathematically correct formal grounds for designing and maintaining modeling business games require adequate mathematical apparatus [1, 2, 3], like the Cartesian closed categories theory [6]. Here it is shown that scalability and "switching" of the semantics in the juridical business games can be attained through the use of the applicative languages and calculation models which agree with the category approach [2]. Topicality of development of the applicative solutions is dictated by the necessity to change the computational paradigm [1,7].

2. Approach to constructing context controlling semantics

While constructing semantics (intentional logic semantics included) the major difficulty is that the meanings of the phrases must depend on the context i.e. a phrase must have different meaning in differing contexts. In fact this circumstance triggered development of the intentional *logic* [8].

The traditional way to account for the context is to bestow an intentional on a phrase and thus define (by a method) a function that defines change of the meaning of a phrase depending on the context. Construction of the efficient intentional *logic* system largely depends on availability of a convenient way to define such a function.

In the majority of known approaches to formalization of the intentional *logic* this issue is not paid much attention to and as a result the function is defined simply as a link between the meanings and the contexts

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

(i.e. presented as a table). Realization of a business game on this approach is rather bulky for it requires presentation of all possible contexts for every phrase. Thus practicality of such systems is limited.

Development of the tools for description of the epistemological level of the business game designing allows for a different approach.

Essentially under this approach the contexts a given certain features which allows to present context as a "bundle" of features. Now when a phrase is put in a context the features that are relevant to the given phrase type are checked out. Depending on the value of the features in a specific context one of the functions (intentional) bestowing the meaning is chosen.

The features of a context can be parameters of this function and in particular directly be included in the extensional (the meaning of a phrase in the context). The suggested way of intentional interpretation is a new one never used in construction of intentional logic before (the authors are not aware of any model formalizations).

Importantly, the necessity to use semantic methods to manipulate juridically meaningful information while designing business games dictates necessity to enhance capabilities of semantic interpretability (or scalability) in the framework of utilization of the formal systems. Yet, the existing formal systems manifest their weaknesses exactly in interpretations manipulations. Though semantics of possible worlds [8] has become a common place in constructing systems of non classic logic all in all the means of defining such worlds and manipulation thereof are underdeveloped (in fact only the means of extracting normal and non normal worlds can be mentioned). Even poorer the situation is with the software realizations.

As is shown above [1, 2, 3] the use of the means of semantic parameterization is substantiated for designing imitational business games in jurisprudence. In this regard the means based on intentional logic are useful. Such include: 1) the means of worlds description including algebraic type techniques; 2) means of construction of evaluating images that define nature of utilization of the worlds and provide content of interpretations in individual semantic components; 3) means of controlling the worlds or correlations and their aggregates.

Suggested approach to representing semantics with control over contexts agrees beautifully with applicative techniques working for objects and constructing computing systems (ACS).

3. Approach to composition and specialization of semantic mechanisms

In terms of conceptual modeling specifics of the practical juridical applications in modeling business games manifests in necessity to maintain hierarchically organized systems of notions based on the volatile classification grounds. Utilization of the foundation as classification of the characteristics that differ in interpretations in different semantic components or environment conditions require their correct correlation in the framework of the model as well as development of the ways of their description and respective specialized semantic mechanisms.

Some objects in terms of the considered approach are formed by composition other are introduced by subject oriented specialization of objects. Essentially in this approach the methods of object composition for the subject matter may be interpreted as applicative objects and be treated in the framework of ABC on the general grounds (including editing, debugging, storing in external memory (e.g. database) etc.). In practical terms they can be represented as specialized combinators, securing subject oriented methods of accounting the context.

Thus the means of controlling semantics for the jurisprudence business games provide correlation between the notions meaningful for modeling a subject matter or establishing subject oriented mechanisms of controlling the same.

Example 1 (application of the conceptual units). Consider an application of conceptual units x (functions) and y (argument), that we name xy. Suppose that the subject oriented context accounting specifics for the chosen subject matter manifests itself in that the x function turns out to be independent of the context z, which must only be accounted for in computation of the argument y. Then the semantic equation reflecting this specificity shall look as

$$[xy]\ z = x\ (yz),$$

and the semantic presentation may be expressed as combinator of the composition

$$\mathbf{B}\ xyz = x\ (yz).$$

Example 2 (modification of context). Suppose now in the conditions of the previous example the context shall be modified using a function φ before any further computations. Then the corresponding presentation can no more be expressed as the combinatior **B** shall require a different combinatior $\mathbf{B}_{\varphi,}$ having a combinatory characteristic

$$\mathbf{B}_{\varphi}\ xyz = x\ (y\ (\varphi\ z)).$$

Using general rules of applicative computing systems (a rule of infinite convolution [2, 4], represent $\mathbf{B}_{\varphi}$ as an application $\mathbf{B}_{\varphi} = \mathbf{F}\varphi$ for an **F.** Then we have

$$\mathbf{F}\ \varphi xyz = x\ (y\ (\varphi\ z))$$

and

$$\mathbf{F}\ \varphi xyz = x\ (y\ (\varphi\ z)) = x\ (\mathbf{B}y\varphi z) = \mathbf{B}x\ (\mathbf{B}y\varphi)\ z$$
$$\mathbf{B}x\ (\mathbf{B}y\varphi)\ z = \mathbf{B}x\ (\mathbf{CB}\varphi y)z$$
$$\mathbf{B}x\ (\mathbf{CB}\varphi y)z = \mathbf{B}\ (\mathbf{B}x)\ (\mathbf{CB}\varphi)\ yz$$
$$\mathbf{B}\ (\mathbf{B}x)\ (\mathbf{CB}\varphi)\ y = \mathbf{CB}\ (\mathbf{CB}\varphi)\ (\mathbf{B}x)\ yz$$
$$\mathbf{CB}\ (\mathbf{CB}\varphi)\ (\mathbf{B}x)\ yz = \mathbf{B}\ (\mathbf{CB}\ (\mathbf{CB}\varphi))\ \mathbf{B}\ xyz$$
$$\mathbf{B}\ (\mathbf{CB}\ (\mathbf{CB}\varphi))\ \mathbf{B}\ xyz = \mathbf{CBB}\ (\mathbf{CB}\ (\mathbf{CB}\varphi))\ xyz$$

Consecutively applying combinatory **B** we get:

$$\mathbf{F}\ \varphi xyz = \mathbf{B}(\mathbf{B}(\mathbf{CBB})(\mathbf{CB}))(\mathbf{CB})\varphi\ xyz,$$

where $\mathbf{C}\ xyz = xzy$. Thus,

$$\mathbf{F} = \mathbf{B}\ (\mathbf{B}\ (\mathbf{CBB})\ (\mathbf{CB}))\ (\mathbf{CB}) \qquad (*).$$

The expression (*) defines the subject oriented method of manipulating a context. Now, the combinatory

$$\mathbf{C}\ xyz = xzy = (xz)y$$

defines dependence of the function on the context and therefore can be considered as a mechanism of specialization of a function in the given computational environment.

The problem domain models providing composition and specialization for the subject oriented mechanisms of inheritance require development of the relevant instruments for maintaining description of subject matters correlated with the engaged theoretical methods. Along these lines the authors have developed a prototype system maintaining a family of the languages of a category type including the semantic elements. The system includes a tunable compiler for a family of languages allowing to construct a model as well as an abstract machine maintaining required calculations. The abstract machine allows calculation of the basic applicative constructions on the basis of the category mechanisms and thus represents the basis of the modeling business game in jurisprudence.

4. Conclusion

Development of the methods and means for problem domain modeling, maintaining description of the subject oriented mechanisms of inheriting and securing their composition and specialization has made it possible to: (1) work out methodology of describing problem domain fragments using subject oriented mechanisms of inheriting for creation of BG; (2) learn potentiality of scaling, composing and specializing of the subject oriented mechanisms in terms of one problem domain by means of ACS; (3) learn potentiality of a) scaling within the framework of one functional basis and b) integration of a number of functional basics and learn methods of their formalized describing and their expressive capabilities.

The submitted solutions in fact represent the means of semantic specialization for applicative models on the category basis and provide theoretical grounds for the interactive procedures of emerging problem domains in ACS and provide a criteria for integration of various applicative models.

References

1. Ismailova L.Y., Kosikov S.V. “Category type languages family providing for the means of semantics control”. In: *International conference Modern IT&(E)learning 2009,* pp. 195–203
2. Zajtsev A.E., Ismailova L.Y., Kosikov S.V. “Applicative models of calculations on the basis of parameterized semantics in the juridical business games”. In: *Proceedings of NRNU MEPhI*, Moscow, 2009, part VI, pp. 16–19
3. Ismailova L.Y., Kosikov S.V. “In quest for lost ontology. Category approach”. In: *Information society technologies – Internet and today society: VII Russian joint conference,* S.-Pb., 2004, pp. 56–58
4. Ismailova L.Y., Kosikov S.V., Zaytsev A.E. “Applicative computational technologies for generating the families of simulating business games”.In: *Proc. of the 11th international workshop CSIT’2009*, 2009, pp. 35–39
5. Wolfengagen V.E. *Methods and means of computation with objects*. Moscow: «Center UrInfoR» Ltd., 2004.
6. Kosikov S.V. *Information systems: category approach*. Moscow: «JurInfoR-Press», 2005.
7. *Applicative computing systems* (Conference ABC’2008 (edited by Ismailova L.Yu.)), Moscow: Institute «JurInfoR-MSU», 2009.
8. Scott D. “Advice on modal logic”. *Semantics of modal and intentional logic,* Moscow, Progress, 1981, pp. 280–317.

Using Neural Networks to Securing Communications Management Systems

Igor Halenár
Faculty of Materials Science
The Slovak Technical University of Technology
Trnava, Slovakia
e-mail: igor.halenar@stuba.sk

Michal Eliáš
Faculty of Materials Science
The Slovak Technical University of Technology
Trnava, Slovakia
e-mail: michal.elias@stuba.sk

Abstract[1]

The common feature of modern automation is opening local media technology facilities towards publicly accessible networks, surveillance on devices and remote management technology. As a result of this, is the penetration of undesirable elements into control and product networks from internet. It is in communications and control networks need to incorporate active agents for providing a secure access to individual components of the technological process. This can be reached by implementation of computer systems providing security control over communications networks. These control systems are usually a dedicated computer systems, connected between internet and local data or production network, and running software on these systems is using conventional methods, like expert systems. We can also to combine conventional methods used in expert systems, in conjunction with methods of neural network technology.
The proposed solution to the control system in this work, defines the way for identification of data elements in the transmission network, solves the transformation parameters for input into neural networks and defines the type and the appropriate neural network architecture.
This is also supported by experiments with different types of architecture and activation functions of neural networks, and then tested in real environments. The result of this is a functional system design, with the possible practical application.

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

1. Introduction

In present days is mostly used a network communication based on TCP / IP protocols. In the case of control systems (PLCs, controllers, actuators, sensors, ...) we are using standard industry protocols based on other methods of communication as a standard in communication in personal networks [1].
The trend is to combine types of industrial ethernet networks and traditional technologies. The results are different technology to post the TCP / IP (Ethernet / IP, Modbus-TCP, PROFINET) that address the compatibility of traditional computing and automation systems, but also bring to the automation network errors and security risks [2].
Therefore, the communication and control networks we need to incorporate active agents to provide data access to individual components of the technological process.

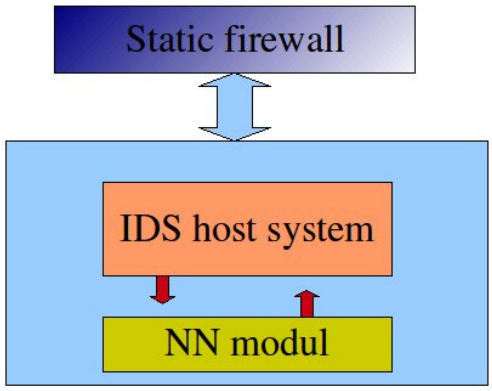
2. System design

The substance of the protection of these important data can be summarized into a few specific points:

1. Ensuring access to data only to authorized persons, or system
2. Guaranteeing the authenticity of information processed
3. The possibility of uniquely identifying a data files creation, the changes, or removal
4. Prevent the possibility of theft of sensitive data
5. Ensuring the availability of data, if necessary

In the design and operation of communications systems with the most emphasis is usually on their functionality. If the nature of the use of information systems, where it is defined by legislation, does not address security aspects, information security is often perceived as an extension of the system and not as part of an information system.
To ensure that communications of systems via a computer network, respectively communication interface, is secure, is needed the whole set of technical and programming resources. Protection is happening

on several levels and is performed by various systems such as firewalls, IDS (Intrusion Detection System), IPS (Intrusion Prevention System) devices or protocols. With joining these elements of security and detection technology with technology of neural networks we can reach a separate intelligent network security systems - intelligent firewall, which contains knowledge about the potential security risks and threats in the computer network and is able to dynamically adapt to the possible failure of a communication system attack, or more complex forms of infiltration (exploits, out of band communications, the analysis of covert communication channels).



The whole proposed system (shown on picture) is suitable to build on existing IDS. Of course the existing system must meet certain requirements:

1. open system - it must be possible to program own extension modules
2. possibility to capture packets - IP and TCP / UDP packets are a base of communication, it should be possible to capture these packets in the communication network
3. registration of packets - an important point because of the possibility of repeating cycles of learning a neural network
4. existence of a communication interface - to manage the active part in data transmission network (firewall)

As a host system is suitable IDS Snort (GNU GPL license) [8].

2.1. Data identification

For validation of correct data transfer is appropriate to begin tracing the values of individual fields in the data structure of TCP and IP protocols, thus easily identify the correctness of packets. Given the large number of protocols used and sub-protocols and their combinations, it is appropriate to choose the parameters governing the selection of the parameters of data packets.

After preliminary analysis of options and the available literature [3], we propose that the characteristics of data packets in this case we specify by the following parameters:

- ID protocol - the protocol type associated with packet
- Source port - Number of TCP / UDP port of the source system
- Destination port - Number of TCP / UDP port of the target system
- Source Address - IP address of the source system
- Destination Address - IP address of the target system
- ICMP type - the type of ICMP packet
- Length of data transferred - the size of the data packet in bytes
- FLAGS - signs in the protocol header
- TCP window size - window size parameter of TCP packet

This data must be properly transformed to be selected as input to neural network. According to available sources we can choose use of feed-forward neural network learning method with back error propagation. This type of neural network provides sufficient flexibility and applicability to a wide range of tasks, where it is possible to use technology to minimize the objective function NN. To minimize the objective function can use several optimization methods commonly used to minimize the numerical mathematics [7]. The commonly used methods are including gradient methods, which disadvantage is the high number of iteration steps. Because of this disadvantage we can use a variety of other, more efficient and faster optimization methods for adaptation of neural networks. In the available literature these methods are represented by the name "quickprop" (based on Newton's method) [5] or other numeric methods (method of variable steps, entropic normalization models, least squares, etc.). Alternatively, you can use modern methods to minimize using genetic algorithms [4].

2.2. Data transformation

The whole process of data transformation is necessary to perform the simplest and fastest. Therefore, we have to implement programs in AWK executable script for Bourne Shell system.
Adaptations have to be carried out in several stages. The first step is extraction of the required values of the entire data package. We can afford to ignore the account records of the communications network service (ARP, RARP communication), and data to us in terms of work unattractive. The next phase is an appropriate representation of some elements.
Specifically, it is the following parameters:

ID_PROTOCOL, where we make the following transformation:

ID_PROTOCOL	Substitution
TCP	6
UDP	17
ICMP	1

TYPE_ICMP, where we make the following transformation:

TYPE_ICMP	Substitution
ECHO	1
REQUEST	2
NULL	0

FLAGS, where we make the following transformation:

FLAGS	Substitution
no_flags	1
RST	2
FIN	3
PSH	4
URG	5
SYN	6
other	7

After processing by the AWK program we receive data represented by the following table (example):

1,1471751309,0,147175134254,0,0,0,44,1
6,62240183148,80,147175134254,2190,8514,4,1500,0
17,19416092,33239,233104778,1234,0,0,1500,0
6,147175130212,2246,19512215120,80,65535,1,40,0
17,13015680151,4262,2242127254,9875,0,0,224,0
1,147175130212,0,10254247189,0,0,0,36,2

Where comma separated fields from first are id protocol, source ip, source port, destination_ip, destination port, size of tcp window, flag, size of packet, icmp type.
For given the expected use of neural networks as classification is sufficient for the distribution patterns of two groups. In this case, it will be sufficient to divide the patterns to correct packets (CLASS = 1) and those which will be subject to further testing (CLASS = 2). Designation of records is necessary for training a neural network. Records of class packets are stored in a separate table.

2.1. Implementation of neural networks

With reference to the used literature [6] is a neural network with one hidden layer (3 layer neural network) and a sufficient number of hidden neurons, capable of simulating each binary or continuous function with desired accuracy. In our case, we assume nine input neurons and two outputs.
With data transformation was selected 5000 models of the ongoing communication in the test environment. Of this number, we have developed a selective choice of four groups containing 200 models for learning neural network and a group of 100, 200 and 3000 models for testing. That was necessary, because of the experimental verification of the results obtained for neural networks with different numbers of neurons in the hidden layer and different trigger functions for hidden and output layer. As the activation function of hidden layer was used in all cases the standard sigmoidal activation function:

$$x_i = f\left(in_i\right) = \frac{1}{1+e^{-cin_i}}$$

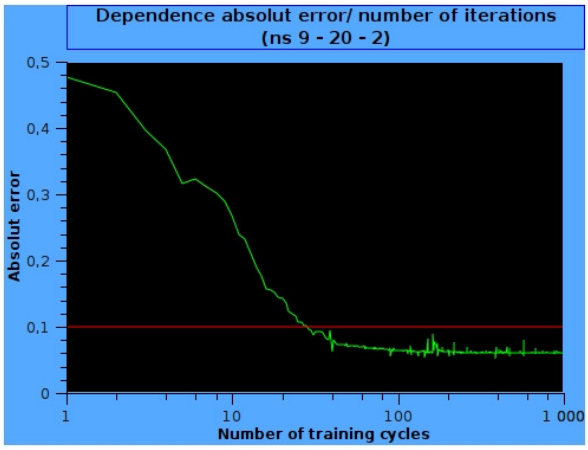
As the activation function of output neurons in the various experiments with neural networks used in addition to sigmoidal a linear activation function:

$$x_i = f\left(in_i\right) = in_i$$

In experimental phase were made several attempts and tests with different numbers of neurons in the hidden layer. Specifically, the network involved in the 9-6-2, 9-10-2 and 9-20-2 (input - hidden - output layer). The same network with different trigger function for the output layer had different results, which resulted in detriment of the use of linear activation function. In the case of application of a linear function for network 9-6-2, the absolute error of neural network after one thousand iterations was 0,23104 . It is a stable value, which is after approximately 200 training cycles not changing. The same network with sigmoidal trigger function achieves the same conditions, the absolute error is lower – 0,17905.
Similar differences are also reflected in another try. It turned out that this proposal neural network generally achieve better results with sigmoidal trigger function for output layer.
As the algorithm to minimize the objective function was chosen quickprop method based on Newton's method [4].

As the results of the experiments we noted, that the best output achieves the network architecture with twenty neurons in hidden layer.
The results obtained after on thousand iterations correspond to the absolute error 0.06041, while the number of patterns in training set, which have achieve the required tolerance, was increased to 96.67%.

Training the network by the increasing number of iteration did not show better results than those in one thousand iterations. Achievements of error in this case meet the required parameters for detection.



By the testing with the test group specimens, network showed good separation of the value of objects. The total number of specimens in the test group was one hundred, the number of samples in the test group to meet the tolerance (0.10) is 98.56%.
From the above results it is clear that the best classification ability of the sample reaches neural network with 20 neurons in the hidden layer. Experiment with applying linear activation function for output layer showed approximately the same variation in the quality of decision-making system as in previous cases.

3. System testing

After training the neural network, values of synaptic weights were stored and we brought to the network current data.
The test set included 10 models of defective packets, particularly in the first instance attempt to communicate to non-standard TCP ports. For the experiment we used a sample size of 3000 data packets. Format of input samples were prepared according to the required parameters. The success of packet error detection was 100%.
The output values of neurons of output layer are listed below:

Packet #	Neuron 3-1	Neuron 3-2
163	0,0000	1,0000
953	0,0000	1,0000
2105	0,0950	0,9050
2191	0,0950	0,9050
2215	0,0950	0,9050
2228	0,0950	0,9100
2345	0,0000	0,9050
2599	0,0300	1,0000
2849	0,0950	1,0000

In the second case the test file pattern contains the same number of errors 10 packets. In this case subject contains the combination of parameters non-standard TCP window size (less than 1028) and erroneously referred to the combination of symptoms of SYN and FIN (FLAGS = 7).
For the experiment was used the same large sample, with size of 3000 data packets. Format of input samples were prepared according to the required parameters. The success of packet error detection in this case has reached 90%, namely mistakenly was recognized by neural network packet with the serial number of the 895th. The results are readable and in the error tolerance (0.1). This shortcoming can be eliminated by better training of the neural network, respectively with better choosing models to training set. The output values of neurons are listed below:

Packet #	Neuron 3-1	Neuron 3-2
464	0,0000	1,0000
895	0,0000	1,0000
1342	0,0950	0,9050
1486	0,0950	0,9050
2215	0,0950	0,9050
2350	0,0950	0,9100
2441	0,0000	0,9050
2559	0,0300	1,0000
2611	0,0950	1,0000
2968	0,1000	0,9100

4. Conclusion

Based on the results obtained, we can say that the implemented system for identifying data packets allows the detection of any data transmitted in the communication network. As seen from the results presented, the type of detected parameters of the transmission of packets depends on the choice of training set.
Designed and trained neural network is able, within the specified tolerances, satisfactorily classify individual data packets. We have shown that the possibility of applications of artificial intelligence technologies in solving problems related to verification of data transfer in management and communications networks.

References

1. Hansen, K.: “Redundancy Ethernet in industrial automation.” In: *Emerging Technologies and Factory Automation, 2005. ETFA 2005. 10th IEEE Conference on Volume 2*. Catania Italy :CRC Press, 2005. s. 947 – 952. ISBN 0-7803-9401-1
2. Trnka A, Tanuška P: Datawarehousing and datamining methodologies as a one of the possible benefit in control process. In: *Aktuaľnye problemy i innovacii v ekonomike, upravlenii, obrazovanii, informacionnych technologijach : materialy meždunarodnoj naučnoj konferencii (12-15 maja 2009 goda, Stavropo=B5-Kislovodsk).* ISSN 2074-1685. - =E8. 5, zv. 3 (2009), s. 84-87.
3. Cannady J.: Neural Networks for Misuse Detection. In: *Proceedings of the 1998 National Information Systems Security Conference (NISSC'98).* Berkeley: USENIX Association, 1999. s. 443–456
4. Schreiber, Peter - Važan, Pavol - Tanuška, Pavol - Moravčík, Oliver: Production Optimization by using of Genetic Algorithms and Simulation Model. In: *DAAAM International Scientific Book.* - ISSN 1726-9687. - 2009. - , 2009. - ISSN 978-3-901509-69-8, s. 169-176
5. Haykin S.: Neural Networks and Learning Machines, third edition. Canada: Prentice Hall, 2009. 936 s. ISBN 01-31471-39-2
6. Kvasnička V. et al: Úvod do teórie neurónových sietí. Bratislava: IRIS, Bratislava, 1997. 140 s. ISBN 80-8878-30-1
7. KEBÍSEK, M. – SCHREIBER, P.: The possibility of utilization of neural networks at the data mining. In: *CO-MAT-TECH 2004: 12th. CO-MAT-TECH. Trnava, Slovak Republic,* 14th-15th October 2004. STU Bratislava, 2004. s. 589-595. ISBN 80-227-2117-4
8. http://www.snort.org/snort/license/gpl

A probabilistic ant colony algorithm for the problem of packing rectangles and circles

A. Valeyeva
Department of Computer Science and Robotics
USATU
Ufa, Russia
e-mail: aida_val2004@mail.ru

R. Fayzrakhmanov
Department of Computer Science and Robotics
USATU
Ufa, Russia
e-mail: rishat_faiz@mail.ru

Abstract[1]

In this article we consider the problem of packing of circles and rectangles in a rectangular area with given sizes (*L,W*). The circular pieces are of *n* different types with radii r_i; *i = 1,…,n*. The rectangle pieces are of *m* different type with length l_j and width w_j, *j=1,…,m*. We solve this problem using a heuristic ant colony algorithm, based on the population. This algorithm uses a modification the well-known procedure which is used for packing circles. We compare the performance of proposed approach to that of existing approximate and exact algorithms on several problem instances taken from the literature. The computational results show that the proposed approach produce high-quality solutions within reasonable computational times.

Keywords: Packing; Heuristic; Optimization, Population, Ant Colony;

1. Introduction

In many industrial sectors, minimizing waste is a critical issue. Our paper studies the circular and rectangle packing problem where a set of circles and rectangles needs to be packing on a rectangular stock sheet of fixed width and length. Its theoretical importance, this problem arises in numerous cases such as the glass, plastic, metal industries. The formulated problem is a case of a known problem of rectangular packing in a strip which in turn is, a case of a NP-difficult problem of one-dimensional packing in containers [1]. Thus the considered problem is NP-hard in strong sense. Consequently, heuristics must be used to solve large-scale real-world problems.

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

Various statements of problems two-dimensional cutting and packings are known. For packing only rectangles or circles many results are received. However, there are not many results for simultaneous packing of circles and rectangles problem. For example, in work [2] the genetic algorithm is used for decision the problem of packing of various difficult figures of the different sizes. Hifi et al. [5] study the problem of cutting a rectangular plate R of dimensions (L;W) into as many circular pieces as possible. In work [6], Rudnev have designed a probabilistic tabu search algorithm for the problem of packing circles and rectangles into a strip. The probabilistic algorithm of search with interdictions using the two-contact coding decision.

In our study, we investigate the circular and rectangle packing problem, where: *L* - length and *W* – width of bin, n_c – many circular pieces, n_r – many rectangle pieces; w_j, l_j – width and length each rectangle pieces, $j \in J_r = \{1,...,n\}$; r_i – radius each circle pieces, $i \in I_c = \{1,...,m\}$. Let's enter rectangular system of co-ordinates *XOY*, at which axis *OX* and *OY* coincide according to the top and left parties of bin. The problem decision is represented in the form of a set of elements <*Xc,Yc,Xr,Yr*>, where <*Xc,Yc*> – vectors of co-ordinates of the centres of circles, (x_{ic}, y_{ic}) – co-ordinates of the centre of i-th circle; < *Xr,Yr*>– vectors of co-ordinates of rectangles, (x_{jr}, y_{jr}) - co-ordinates of the top left corner of a rectangle accordingly on an axis *X* and *Y*. The set of elements <*Xc,Yc,Xr,Yr*> is called as admissible packing if following conditions are satisfied:

1. The parties of rectangles are parallel to sides of rectangular stock plate (an orthogonality condition);
2. Rectangles and circles do not overlap each other;
4. Rectangles and circles do not overlap sides of bin;

At performance of conditions of admissibility it is required to maximize the usage of bin.

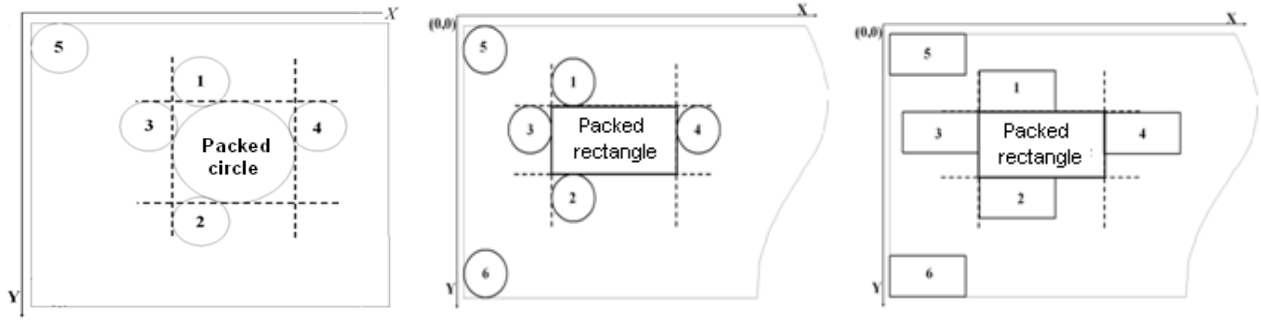
In this paper, we present a constructive ant colony algorithm for rectangles and circles packing problem. The algorithm uses a modified version of the adapted

best local position (*ABLP*) approach [5] for a dense deployment circles and rectangles.

2. An approach for circles and rectangles packing problem

The modified adapted Best Local Position procedure, namely *ABLP+,* constructs a "good" layout of an ordered set of circular and rectangle pieces.
Given an ordered set of pieces, the ABLP+ proceeds as follows. It starts by placing the reference point of the first piece on the upper left-most position of the rectangular stock sheet R; i.e., on the *(0; 0)* coordinates point. It places the second piece on the upper left-most available position avoiding its overlap with the first piece and respecting the width constraint. In general, the constructive approach chooses for each piece P_k to be packed the best (upper left-most) position among a set of possible positions. This set is constructed as follows:

First, each already placed piece P_i, with reference point $(xi; y_i)$, generates a set of four positions. The first and second positions are simultaneously adjacent to P_i and to the horizontal line $y = y_i$. The third and fourth positions are simultaneously adjacent to P_i and to the vertical line $x = x_i$. These four positions along with the corner of the rectangular stock sheet are illustrated in Fig. 1.
Second, each pair of adjacent circles C_{i1} and C_{i2} generates two possible positions. These positions, illustrated in Fig. 2, make C_i simultaneously adjacent to C_{i1} and C_{i2}.
Third, the corners of bin generate additional positions.
Fourth and last, the resulting set is reduced by eliminating duplicate positions, cause P_i to overlap an already placed piece, and positions that violate. Evidently, the resulting set differs according to the piece P_i to be packed.



a) b) c)

Figure 1: Positions at placing of two piece

3. An constructive ant colony algorithm based on population.

For the approximate solution packing problem of rectangles and circles developed a constructive ant colony algorithm using the general scheme of *P-ACO*, described in [7]. The founder of the ant colony algorithm is M. Dorigo. This algorithm simulates the behaviour of real ant colonies, in the course of experiments on which it was found that the ants find the shortest path from the nest to the source by olfaction. At the same time while moving ant "marks" the traversed path drops by fragrant liquid - pheromone that attracts other ants. Faced with pheromone trail, the ant is more likely to be followed on it, causing a pheromone, from what should be enhanced. Are known various modifications of the algorithm, the best of them in solving combinatorial optimization problems recognized Ant Colony Optimization (*ACO*), Ant Colony System (*ACS*) and the Max-Min Ant System (*MMAS*). *P-ACO* algorithm based on *ACO* algorithm and on idea of genetic algorithm.
To apply the ant colony algorithm for combinatorial optimization problems is necessary to determine its main characteristics: the component solutions, the heuristic information (number, independent of the solutions found in previous iterations, and reflects the desirability of adopting one or another component of the decision); rule changes pheromone (a number that indicates how often the agents used components of the solution at the previous iteration). The solution, built by each agent, is formed from the components based on the accumulation and use of statistical information - artificial pheromone traces and task-specific heuristic information.
Consider the general scheme of the algorithm *P-ACO*:
Input: L - length and W – width of bin, n_c – many circular pieces, n_r – many rectangle pieces; w_j, l_j – width and length each rectangle pieces, $j \in J_r = \{1,...,n\}$; r_i– radius each circle pieces, $i \in I_c = \{1,...,m\}$.
Exit: the best found decision (a packing map)
1. Initialization of parameters of algorithm: k – size of population; m – count of agents; α – factor of influence of the pheromone; β – factor of influence of the heuristic information; τ_{init} – initial value of pheromone; τ_{max}– maximum value pheromone.
2. Repeat until stopping criterion is satisfied

2.1. To choose a component from set $I_c \cup I_r$ and to place it in the top left corner of a bin

2.2. For each of m the agents who have not finished construction of decisions to execute:

2.2.1.To choose the following component from set $I_c \cup I_r$

2.2.2. To add the chosen component to partially constructed decision by means of procedure *ABLP+*

2.3. For decisions constructed by agents define the best decision

2.4. If |*P*| = *k*, to remove the decision from population with a strategy

2.4.1. To execute updating pheromone.

2.5. To add the best decision in population *P*

2.5.1. To execute updating of pheromone

3. To give out result – the best decision

Each agent builds a solution, starting with some initial component, and at each step adds a component of the solution according to some probabilistic rule:

$$p = \frac{\tau_{ij}^{\alpha} \cdot \eta_{ij}^{\beta}}{\sum_{k \in I_c \cup I_r} \tau_{ik}^{\alpha} \cdot \eta_{ik}^{\beta}} \qquad (1)$$

where, τ_{ij} - significance level of pheromone for the transition from state *i* to state *j*, η_{ij} - value of heuristic information for the piece to be packed (the area piece).

At the beginning of the algorithm *P-ACO* population is empty. During the first k iterations of the population *P* is added to the best solution obtained by agents at each iteration. Since *k+1*-th iteration, the best solution for this iteration is a candidate for addition to the population. When a population is formed, | *F* | = *k*, you need to update it.

Typically, the following strategy for updating the population: Age (removed the most "old" solution); Quality (the worst solution is removed); Prob (removing solutions from the population occurs with some probability).

Update pheromone is only for making the population: if the solution is added to the population, to the value of pheromone is added value, if the solution is removed from the population, the value of pheromone is reduced by the

4. Computational results

Were carried out numerical experiments on a computer Intel Quadra 2,4 GHz. Solved the problem of packing circles and simultaneously packing circles and rectangles in the strip for randomly generated examples, as well as examples of the international digital library. This involves the analysis of the quality of the algorithm by comparison with the lower and upper bounds. In addition, the results were compared with those obtained by other authors.

Experiment number 1. The circles packing problem. Test examples are taken from article [6]. The calculations were performed with the following parameters: $\alpha = 1,8$; $\beta = 3,9$; number of agents $m = 12$; population size $k = 15$; initial value of pheromone = 0,05; maximum value of pheromone = 0,85.

Table 1 shows the following characteristics: *W* –width of strip, n_c - many circular pieces; *LB* - lower bound on the optimal length of the strip and presented the solution of packing circles obtained by genetic algorithm CAGA, the method branches and boundaries SY, greedy heuristic procedure Q1.5 and the algorithm of probabilistic tabu search TABU SEARCH (TS). Time of the algorithm CAGA was 30 minutes on a computer with a processor Pentium III 733MHz. Hours branch and bound method is not specified, except for an example SY6, for which it took one hour IBM PC / AT 486. The average running time Q1.5 was approximately 40 minutes for example SY1-SY4 and 18 hours for example SY5 - SY6 and on a machine with a processor Athlon XP2000+. Hours of probabilistic tabu search algorithm and the P-ACO was 30 min. In the last column of Table 1 shows the results of found by Rudnev A.S. using the commercial package GAMS (http://www.gams.com/). Solving them was obtained by connecting to the GAMS decisive nucleus BARON. For each example, the package has been allocated 20 hours, during which it was possible to find solutions only for the examples SY1-SY4. The algorithm of the P-ACO has found new record solutions to the examples SY2, SY4 and SY5 (see Table 1). For the first three the best solution is found the algorithm TS, for the sixth example of the best solution is found the algorithm Q1.5.

Table 1: The length of the strip for the circles packing problem

Inst	***W***	n_c	***LB***	**CAGA**	**S.Y.**	**B1.5**	**TS**	***P-ACO***	**GAMS**
SY1	9.5	30	14.550	18.368	17.491	17.291	**17.262**	17.340	19.850
SY2	8.5	20	12.160	15.239	14.895	14.535	14.515	**14.182**	17.721
SY3	9	25	12.234	15.413	14.930	14.470	**14.430**	14.480	16.758
SY4	11	35	19.907	24.979	24.355	23.555	23.502	**23.245**	25,887
SY5	15	100	31.282	38.851	38.047	36.327	36.707	**35.865**	----------
SY6	19	100	31.784	39.650	38.647	**36.857**	37.422	36.958	----------

Experiment number 2. For this experiment, test samples were taken from article [6]. Ant colony algorithm is executed with the same parameters as the previous experiment. Test classes were generated

randomly so that each sample was already known some upper bound of its optimal solution. Generation of examples in carried out as follows. The rectangular area the size of $W * UB$, where UB - upper bound given the optimal length of the strip, W - width of the strip, divided into squares n_c and nr rectangles waste-free manner. Then, each square was replaced by the circles of a diameter equal to the side of the square. In total 18 examples were generated with the number of subjects from 8 to 60. During the experiment, *P-ACO* algorithm was run 10 times for each sample. Consideration was oriented case of a problem, i.e. corners of rectangles by 90 degrees were prohibited.

Table 2 shows the parameters of an example: W-width of strip; n-number of subjects; n_c – number of circles; n_r - number of rectangles; LB and UB - lower and upper bounds, algorithms for TS, P-ACO shows the results of algorithms.

Numerical experiments show that this problem developed algorithm in a reasonable time of receiving the decision, which is close to the lower estimates, as well as when compared with the algorithm TS, P-ACO algorithm has shown better results in the following classes of problems: CR3P02; CR5P01; CR5P02; CR5P03; CR6P03. At the other classes of problems the best solution showed tabu search algorithm. In the last column of Table 2 shows the results only for the examples CR1P01-CR1P03, obtained by using the package GAMS.

Table 2: The length of the strip for the circles and rectangles packing problem

Inst	Parametrs						The heuristic		
	W	***n***	n_c	n_r	***LB***	***LU***	**TS**	***P-ACO***	**GAMS**
CR1P01	10	7	3	4	8.986	10	10	10	10
CR1P02	10	8	4	4	8.960	10	10	10	10
CR1P03	10	7	4	3	8.689	10	10	10	10
CR2P01	20	17	7	10	17.833	20	20	20	----------
CR2P02	20	17	8	9	17.317	20	19,909	19.980	----------
CR2P03	20	17	6	11	17.693	20	20	20.34	----------
CR3P01	40	25	10	15	13.460	15	15.266	15.345	----------
CR3P02	40	25	6	19	13.830	15	16,170	15.980	----------
CR3P03	40	25	14	11	13.181	15	15	15.241	----------
CR4P01	60	29	10	19	27,611	30	30,972	32,682	----------
CR4P02	60	29	14	15	26,212	30	30,430	34.721	----------
CR4P03	60	29	6	23	28,455	30	31,000	28,840	----------
CR5P01	60	49	20	29	53,143	60	63,040	61.898	----------
CR5P02	60	49	16	33	53,962	60	63,209	62.465	----------
CR5P03	60	49	12	37	56,995	60	63,580	58,762	----------
CR6P01	60	73	39	34	78,297	90	93,147	94.907	----------
CR6P02	60	73	35	38	78,762	90	92,934	91.432	----------
CR6P03	60	73	31	42	78,977	90	93,982	92.840	----------

This algorithm was tested on real data, the company "Thermophysics". This company is engaged in the manufacture of nake-to-order. Some of the products of this type of material are combustion chamber, central burner and etc. To produce a sheet of width W = 1m, length L = 2 m, thickness H = 1.5 mm. from different types of metals (stainless steel, etc.). For example for the manufacture of the combustion chamber, central burner using such blanks as: disk, membrane, diaphragm, wall, cone, cylinder and etc.

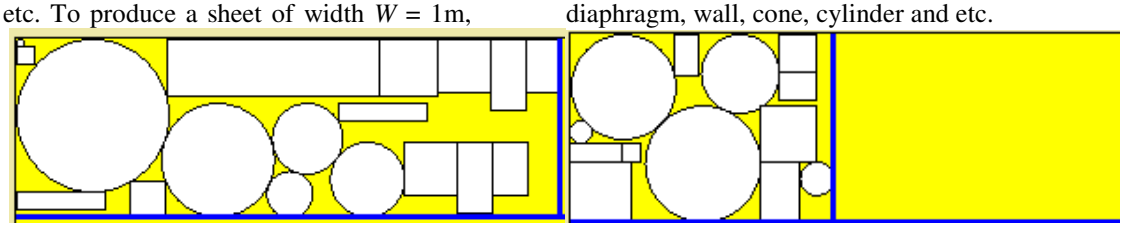


Figure 2: Illustration a packing map

Various experiments showed that in the enterprise application of the proposed algorithm gives a saving of material within 6-9%.
This problem also arises in other kind of industrial when you want an industrial material is cut into rectangular and circular blanks for use in the manufacture of final products (e.g.: round and rectangular mirrors, tables and manufacturing of metal constructions, etc.).

4. Conclusion

We have solved the rectangles and circles packing problem using Constructive Ant Colony algorithm and modified constructive procedure *ABLP+*. Testing of different problem taken from the literature shows that the constructive heuristic yields good solutions within a short computing time. The proposed algorithm finds new record for some test examples. This al Later will apply the developed algorithm for three-dimensional packing problems.

Acknowledgement

The research is supported by grant of President of Russian Federation № 65497.2010.9 for Leading Scientific Schools.

References

1. Dorigo M., Di Caro G., Gambardella L. "Ant Algorithms for Discrete Optimization" *Artif.Life*. 1999. Vol.5. No.3. pp.137-172.
2. Burke E., Kendall G. "Applying Ant Algorithms and the No Fit Polygon to the Nesting Problem" Accepted for the 1999 International Conference on Artificial Intelligence, Monte Carlo resort. Las Vegas. Nevada. USA. 1999.
3. Di Caro G., Dorigo M. Extending "AntNet for best-effort Quality-of-Service routing." Unpublished presentation at ANTS'98 – From Ant Colonies to Artifical Ants: First International Workshop on Ant Colony Optimization, October 15-16 1998.
4. Valeeva A.F. "Constructive methods for orthogonal packings and cutting problems". Ufa,2006.
5. Hifi M., M'Hallah R. "Approximate algorithms for constrained circular cutting problems". Computers and Operations Research, 2004, N31, P. 675-694.
6. Rudnev A.S. "Probabilistic tabu search algorithm for the problem of packing circles and rectangles into a strip". Discrete Analysis and Operations Research. 2009. - T. 16, No. 4. - S. 61 - 86.
7. M. Guntsch, M. Middendorf. "Applying Population Based ACO to Dynamic Optimization Problems" Proc. 3rd Int. Workshop (ANTS 2002), Lecture Notes in Computer Science, vol., 2463, Springer, Berlin, p.p. 11-122

Computing the KPZ Equation Using GPU Acceleration

Henrik Schulz
Forschungszentrum Dresden-Rossendorf e.V.
Department of Information Technology
Dresden, Germany
e-mail: h.schulz@fzd.de

Géza Ódor
Research Institute for Technical Physics and Materials Science
Complex Systems Group
Budapest, Hungary
e-mail: odor@mfa.kfki.hu

Jeffrey Kelling, Karl-Heinz Heinig, Bartosz Liedke
Forschungszentrum Dresden-Rossendorf e.V.
Institute of Ion Beam Physics and Materials Research, Division of Theory
Dresden, Germany
e-mail: j.kelling@fzd.de, k.h.heinig@fzd.de, b.liedke@fzd.de

Nils Schmeißer
Forschungszentrum Dresden-Rossendorf e.V.
Department of Information Technology
Dresden, Germany
e-mail: n.schmeisser@fzd.de

Abstract[1]

In this talk we present an approach for fast parallel computation of the Kardar-Parisi-Zhang equation (KPZ). Acceleration of these computations is always a major task since the run time of the simulations representing sequential algorithms is in the range of several months. We developed an implementation for the KPZ equation on graphics processing units (GPU) using the Compute Unified Device Architecture (CUDA). Experimental results show a significant speed-up compared to implementations on conventional processors.

1. Introduction

Graphics processing units are a special kind of processors, originally developed to perform computation for computer graphics. These operations have to run very fast and this is achievable by high parallelism of the processing units. In 2007 Nvidia released version 1.0 of the CUDA software development kit to use GPUs for general-purpose computation, too. Despite many improvements to GPUs the functionality is still limited. Especially memory operations are a bottleneck due to the limited bandwidth of the bus systems.

Today's GPU devices consist of a number N of multiprocessors, each equipped with M processor cores, shared memory and caches (see Fig.1). Since the total number of processor cores inside a GPU (N*M can be up to 240) is much greater than that of today's CPUs, the major task is to decompose the computational problem into many small fragments. On the other hand there is only a small amount of memory per core (16 kBytes shared over 8 cores). This leads to the problem that memory management is the most important factor for the efficiency of the parallel program.

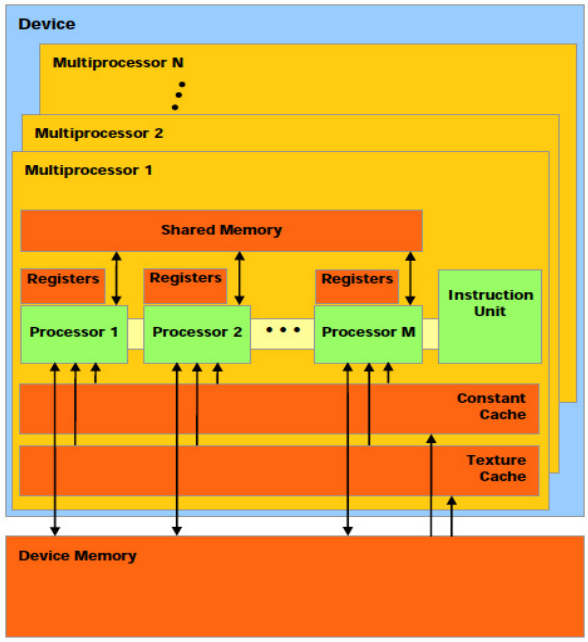


Figure 1: Architecture of Nvidia GPUs [1].

In statistical physics the simulated spaces can often be decomposed into small cells of points which contain

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

dependencies, i.e. the state at a certain point can only be computed considering neighboring points. Inside a cell the necessary data is available in local memory, but between the cells the data has to be exchanged. Since there is no mechanism to communicate between GPU cores directly, one has to use the device memory to exchange data, which is a classical SMP (symmetric multiprocessing) approach.
In this talk we explain the implementation of the KPZ equation on GPU devices (Section 3) and we give experimental results to show the speed-up of this implementation (Section 4). But before we can go on, we need some basic facts about the physics in Section 2.

2. The Roof-Top Model in Statistical Physics

One of the simplest nonlinear stochastic differential equation set up by Kardar, Parisi, and Zhang (KPZ) [2] describes the dynamics of growth processes in the thermodynamic limit. It specifies the evolution of the height function $h(x,t)$ in the d –dimensional space.

$$\partial_t h(x,t) = v + \sigma \nabla^2 h(x,t) + \lambda(\nabla h(x,t))^2 + \eta(x,t)$$

Here v and λ are the amplitudes of the mean and local growth velocity, respectively, σ is a smoothing surface tension coefficient, and η roughens the surface by a zero-average Gaussian noise field exhibiting the variance $\langle h(x,t), h(x',t')\rangle = 2D\delta^d(x-x')(t-t')$. The notation D is used for the noise amplitude and $\langle\rangle$ means the distribution average.

Atomistic models following KPZ dynamics were introduced for exploring surface phenomena in the very early times (for a review see [3]). In one dimension it turned out that a very simple restricted solid-on solid model (see Fig.2) exhibits KPZ surface scaling behavior asymptotically.

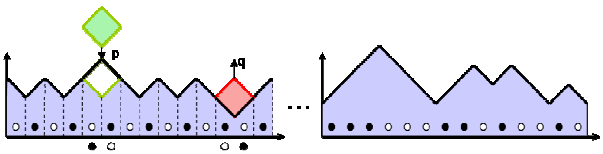


Figure 2. Mapping of the (1+1)-dimensional surface growth onto the 1D ASEP model

In this discrete so-called "roof-top" model the heights are quantized and the local derivatives can take the values $\Delta h = \pm 1$. By considering the up derivatives $\Delta h = 1$ as particles and the down ones as holes, the roughening dynamics can be mapped onto an anisotropic diffusive system of particles (ASEP model) with single site occupancy. The ASEP in statistical physics is one of the most fundamental, well-known nonequilibrium system (for a review see [5]).

The roof-top construction has been extended to 2+1 dimensions [4] by the introduction of octahedra having four slopes. The up edges in the x or y directions can be represented by "$+1$," while the down ones by "-1," and a surface element update is a generalized Kawasaki exchange [6]. The translation of up edges to particles and the down ones to holes of the base lattice maps particle deposition or removal processes onto two simultaneous particle moves one in the x and one in the y direction . One can also consider it as a dimer move in the bisectrix directions of x and y see Fig. below. Therefore, the 2+1 surface dynamics can be mapped onto a "two-dimensional ASEP" of oriented dimers exhibiting hard-core exclusion. The asymmetric drift corresponds to an evolving surface exhibiting KPZ scaling, while the symmetric dimer diffusion is related to the Edwards-Wilkinson behavior [7].

This mapping construction can be extended to higher dimensions and for describing other surface reactions (Mullins diffusion, Kuramoto-Sivashinsky...)[9]. Extensive numerical simulations were done to confirm that these lattice gas models exhibit KPZ scaling in $d=1, 2, 3, 4, 5$ [8], and the surface pattern formation and scaling of Kuramoto-Sivashinsky model was studied [9].

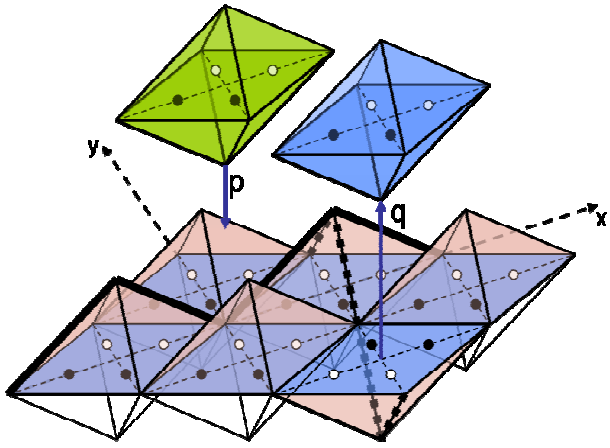


Figure 3. Mapping of the (2+1)-dimensional surface growth onto the 2D particle model

3 Implementation of the KPZ Equation Using GPUs

Our simulation uses the 2+1-dimensional "roof-top" model. Since the statistics of the process depends on random selection of lattice points we can only use domain decomposition for the parallelization. The complete system is located in the main memory and small parts of it are transferred into the graphics memory for the computation. During this transfer the algorithm performs the decomposition of the system into subsystems. Each update needs to consider next neighbors of each lattice point and therefore the boundaries of the subsystems have to be omitted. This leads to errors, but random translation of the

decomposition boundaries after each update is used to minimize the errors (see Fig.4). Random translation means that the complete system is decomposed again with the origin moved in x- and y-direction to ensure that most of the old boundary cells are inside the new subsystems to be part of the computation.

In the kernel now each thread contains its own subsystem which is located in the main graphics memory, since the fast shared memory of the graphics processing cores is much too small. Necessary random numbers are produced by a linear congruent random number generator [10] during computation and its state is stored in the shared memory. This approach is much faster than transferring random numbers from the global memory.

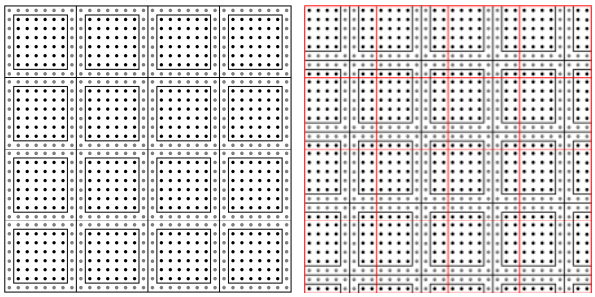


Figure 4: Initial decomposition of the domain (left) and decomposition of the domain after moving the origin (right). Red lines illustrate the new decomposition on the old but moved system.

4 Experimental Results

We tested our implementation on a PC system with two AMD Opteron 2346 quad-core CPUs (1,8 GHz) and one Nvidia Quadro FX 3700 graphics card. The graphics processor consists of N*M=112 scalar processors with a clock rate of 500 MHz. The memory clock rate is 800 MHz. The GPU implementation runs by a factor of 24 faster than a comparable implementation on one CPU core.

The same tests on one card of an Nvidia Tesla S1070 resulted in a speed-up of a factor 41 which is mainly caused by the higher number of scalar processors (N*M=240) and the higher processor clock rate of 602 MHz.

Table 1: Summary of the experimentally discovered speed-up values on different architectures

System	Number of cores used	Processor clock rate	Speed-up
AMD Opteron 2346	1	1,8 GHz	1
Nvidia Quadro FX 3700	112	500 MHz	24
Nvidia Tesla S1070	240	602 MHz	41

5 Conclusion and Future Work

The implementation of the KPZ equation has shown that a significant speed-up for this kind of simulation in the field of statistical physics is possible with the architecture of today's graphics processors. The most important insight of our tests is the fact that the bottleneck of the simulations on GPUs is access to the global memory since random access to lattice points is needed in each step and the graphics memory is not large enough to store the complete system. To overcome this problem we consider using bits instead of bytes to encode the lattice. With this method it is possible to fit the subsystems into the fast shared memory of the GPU cores, but using less threads which leads to a lower computing performance.

Future work will focus on the scaling of this implementation on more than one GPU using GPU clusters and MPI. It turned out that random number generation is important for the correctness and efficiency of the simulation and therefore it will be necessary to compare our results with the results using different generators, for example the Mersenne Twister [11].

Acknowledgement
This project is sponsored by P-MÖB/854 and Hung. Res. Fund OTKA T77629, FP7 OSIRIS and by the DAAD, Project-ID 50450744. Furthermore the authors are pleased to thank Nvidia for supporting the project with high-performance graphics cards.

References

1. Nvidia CUDA Programming Guide.
2. M. Kardar, G. Parisi, and Y. C. Zhang, Phys. Rev. Lett. 56, 889 (1986).
3. A. L. Barabási and H. E. Stanley, Fractal Concepts in Surface Growth Cambridge University Press, Cambridge, UK, (1995).
4. G. Ódor, B. Liedke, and K.-H. Heinig, Phys. Rev. E 79, 021125 (2009).
5. G. Ódor, Universality in Nonequilibrium Lattice Systems, World Scientific, Singapore, (2008).
6. K. Kawasaki, Kinetics of Ising models, In C. Domb and J. L. Lebowitz (eds.) Phase Transitions and Critical Phenomena, Vol.2, chap. 11, p. 443, Academic Press, London (1972).
7. S. F. Edwards and D. R. Wilkinson, Proc. R. Soc. 381, 17 (1982).
8. G. Ódor, B. Liedke, and K.-H. Heinig, Phys. Rev. E 81, 031112 (2010).
9. G. Ódor, B. Liedke, and K.-H. Heinig, Phys. Rev. E 81, 051114 (2010)
10. D. E. Knuth, The Art of Computer Programming, Vol.2, Addison-Wesley, (1969).
11. M. Matsumoto, T. Nishimura, ACM Trans. on Modeling and Computer Simulation Vol. 8, No. 1, January pp.3-30 (1998).

A Study of Odd-Even and Rank Parallel Sorting Algorithms for GPU

Fiaz Gul Khan
Politecnico di Torino, Italy
e-mail: fiaz.khan@polito.it

Omar Usman Khan
Politecnico di Torino, Italy
e-mail: omar.khan@polito.it

Bartolomeo Montrucchio
Politecnico di Torino, Italy
e-mail: bartolomeo.montrucchio@polito.it

Abstract[1]

Sorting algorithms have been studied extensively since past three decades. Their uses are found in many applications including real-time systems, operating systems, and discrete event simulations. In most cases, the efficiency of an application itself depends on usage of a sorting algorithm. Lately, the usage of graphic cards for general purpose computing has again revisited sorting algorithms. In this paper we are presenting an analysis of parallel odd-even and rank-sort algorithms. Their performance for various queue sizes is measured with respect to sorting time and rate. The algorithms have been written to exploit task parallelism model as available on multi-core GPU's using the OpenCL(Open Computing language) specification. Our findings report maximum 5x speed-up of odd-even against rank-sort technique for small queues.

1. Introduction

Parallelism at chip-level is a driving force behind recent advancements in micro-processor architectures. Multi-core CPU's are commonly available in the market and the number of available core's per CPU is expected to increase over the years. A graphics card or graphics processor unit (GPU) is also a common hardware peripheral component that is present in most CPU's for purpose of running advanced graphics based applications. The GPU's on-board processor having upto 240 cores (NVIDIA G80 family specification) [3] can also be used as a co-processor to the CPU, augmenting the idea of parallelism. Current GPU's are capable of out-classing high-end CPU's by a very big margin if compared on basis of GFLOPS, as an example of which NVIDIA Tesla C1060 GPU has peak performance of 933 GFLOP as compared to 107 GFLOP on Intel Core I7 for single precision calculations [4].

Sorting algorithms are critical to many areas of computer science and beyond. It's importance has lead to introduction of many of it's variants, for exploiting both sequential and parallel architectures. In most cases, the efficiency of a program itself depends on selection of an optimum sorting algorithm. Although sequential and parallel sorting algorithms is a main topic in studies of algorithmics since over three-decades, it's performance evaluation on GPU's is still new.

A GPU is a multi-core processor specialized for graphic acceleration and can support thousands of threads running concurrently. Programming on the GPU has been rampant with usage of OpenGL specification which is used specifically for graphic purposes. But recently, the GPU has also been used for general purpose computing purposes, in which case, applications which require heavy number crunching can scale very well on GPU devices. However, it is a very cumbersome task to represent general purpose algorithms in terms of multidimensional operations.

The GPU device used for meeting objectives of this paper is the NVIDIA GeForce GTX-260 which supports 27 microprocessors having 8 cores each, totaling to 216 cores arranged as a streaming multi-processors array[3]. The results have been also compared with another device; the NVIDIA GeForce GT 320 M, having 3 micro-processors with 8 cores each, totaling to 24 cores. Table-I gives an overview of technical data of both these GPU's. The GT 320 M is

[1]**Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

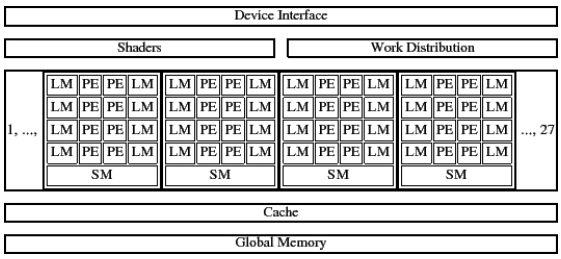


Figure 1: NVIDIA GTX-260 Device Architecture (LM: Local Memory; PE: Processing Element; SM: Shared Memory)

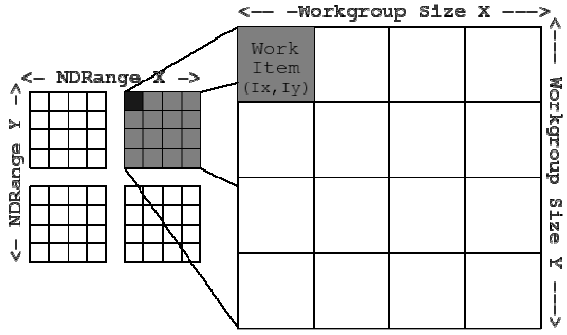


Figure 2: Work-Items (I_n) & Work-Groups (G_n) in two dimensions

more ideally suitable for laptops and hence has fewer numbers of cores for balancing battery power. The more powerful GTX 260 of G80 series scales well for desktop PC's and can require 182W of power and reach up to 105°C [3] at peak performance. These cores are capable of running thousands of threads concurrently. Management of threads is done explicitly in hardware and hence thread context switching time is close to none. To exploit such a large number of threads, the GPU utilizes the single instruction multiple thread (SIMT) architecture model whose main feature is running of data parallel codes. Table-I also shows that CPU-GPU communication is slower than the interaction between threads within any GPU.

Programming for the purpose of general purpose computing on GPU's is possible through CUDA [2], Fire Stream [1] and OpenCL [5][5]. While CUDA targets GPU devices specifically, the OpenCL targets all devices which conform to it's specification. This may include GPU, CPU, and other devices but is still largely focused on the GPU. The GPU has been used for general purpose computing widely both for research as well as in industry. It is particularly popular for modeling of physical sciences, e.g., molecular dynamics [8], fluid dynamics [12], graph algorithms [10] and database queries [11]. In all cases, the types of application involve large-scale heavy computationally intensive operations which can be ported onto a SIMD architecture. GPU's are very optimized for SIMD architectures.

Table 1:Technical Data of GPU's Used

GTX260	GT320M	
216	24	Number of Cores
1242 Mhz	1100 Mhz	Clock Rate
2138 Mb/s	2499 Mb/s	Host - Dev Mem Bandwidth
1618 Mb/s	2777 Mb/s	Dev - Host Mem Bandwidth
93558 Mb/s	10189 Mb/s	Dev - Dev Mem Bandwidth

OpenCL is an extension of C99 language with a few extra keywords and slightly modified thread syntax for running kernel applications. The programming model for OpenCL includes a host program that is able to launch parallel programs called kernels on the target platform/device. Typically the host runs on a CPU while the kernel runs on the GPU. Based on the SIMD/SIMT model, the kernel runs a sequential program across multiple threads. The programmer organizes the problem domain into several work-items, which together make up a work-group. A single program may contain one or more of these work-groups. The collective problem domain is referred to as the NDRange, which can support up to three dimensions. Each work-item within the complete problem domain is identified by an addressing scheme based on the sizes of NDRange and work-groups. The addressing scheme is outlined in Figure-2 for a 2-Dimensional problem. Each unit of work can reference it's work-items local_id, which is the internal addressing scheme within a work-group, it's group_id, which is the addressing scheme for identifying the work-group it is operating in, and the global_id, which is the global address determined by:

$$global_{id} = (group_{id} \times group_{size}) + local_{id}$$

.

The hardware limits for these ID's are set to 512x512x64 for both GTX260 and GT320M GPUs. This essentially means up to 16,777,216 concurrent threads for a three-dimensional problem. Most of these threads are executed in thread blocks, known as a warp. The minimum size of warp on these devices is 32 threads.

We are presenting an analysis of parallel sorting algorithms for many core GPU devices using the OpenCL specification. Our focus is on the odd-even, and rank parallel sorting algorithms. Our results will suggest usage of appropriate algorithm for different areas and applications where sorting is crucial to system performance.

2. Parallel Sorting Algorithms

Sorting on GPU's require transferring data from main memory to on-board GPU global memory. Although

on-device bandwidth is up to range of 140Gb/s, only those sorting techniques are efficient which require minimum amount of synchronization because the PCI bandwidth is to the range of 2.5Gb/s. I.e., synchronization and memory transfers between CPU and GPU will affect system performance adversely. Compared to serial sorting algorithms, parallel algorithms are designed to show more data independence between various elements. Those techniques which involve heavy data dependency are usually sequential sorting algorithms.

2.1. Odd-Even Sort

The odd-even sort is a parallel sorting algorithm and is based on bubble-sort technique. Adjacent pairs of items in an array are exchanged if they are found to be out of order. What makes the technique distinct from bubble-sort is the technique of working on disjointed pairs, i.e., by using alternating pairs of odd-even and even-odd elements of the array. The technique works in multiple passes on a queue Q of size N. In each pass, elements at odd-numbered positions perform a comparison check based on bubble-sort, after which elements at even-numbered positions do the same. The maximum number of iterations or passes for odd-even sort is N/2. The total running time for this technique is $O(\log^2 N)$. The working of the algorithm is as such:

Algorithm 1 Odd-Even Sort

```
1  For k = 1 to N/2 do
2     do parallel
3       if i > i+1 ∀ i%2! = 0
4         swap i, i+1
5     end parallel
6     do parallel
7       if i > i+1 ∀ i%2 == 0
8         swap i, i+1
9     end parallel
10 end for
```

2.2. Rank Sort

There are two phases of the rank-sort algorithm. In the first phase, for each element in queue *Q* of size *N*, the total number of elements less than itself is maintained in another data structure of same size *N*. This is calledthe ranking phase and is depicted in Algorithm-2:

Algorithm 2 Ranking Phase for Rank sort

```
1  For k = 1 to N do
2     do parallel
3      Q[n] = 0;
4      for i = 1 to N do
5        for j = 1 to N do
6          if Q[j] < Q[i] r[i]++
7          else r[j]++
8        end for
9      end for
10   end parallel
11 end for
```

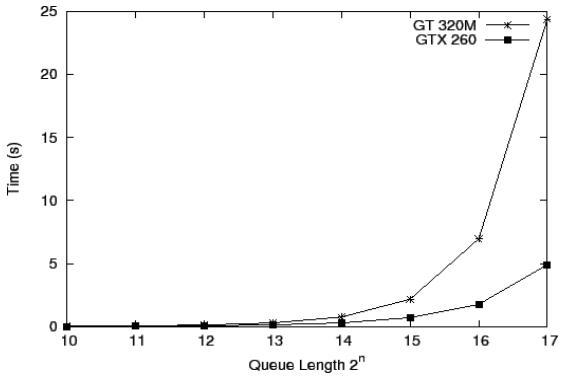
Since each element *n* is compared against n-1 other elements, therefore there are a total of n(n-1) total computational steps. But since the comparison requires sharing of data and not changing of data, the comparison can be made in O(N) total steps for N processors. This also means that the technique is feasible for shared memory architectures.

The second phase involves sorting of elements in queue *Q* based on it's rank. The phase is shown in Algorithm-3:

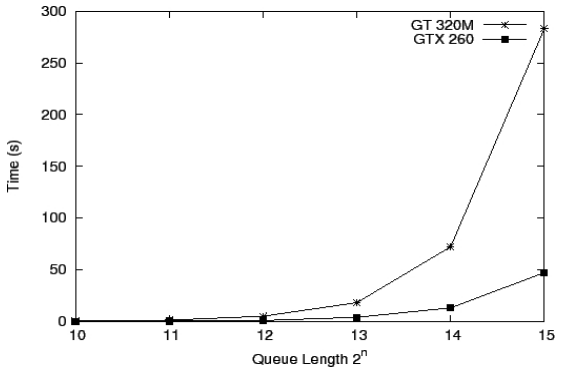
Algorithm 3 Sorting Phase for Rank Sort

```
do parallel
    for k = 1 to N do
      u[r[k]] = a[k]
    end for
  end parallel
  do parallel
    for k = 1 to N do
      a[k] = u[k]
    end for
  end parallel
```

The second phase sorting can be performed in $O(\log_2 n)$ steps. For optimization, the number of elements is divided based on number of processors using m=n/p.



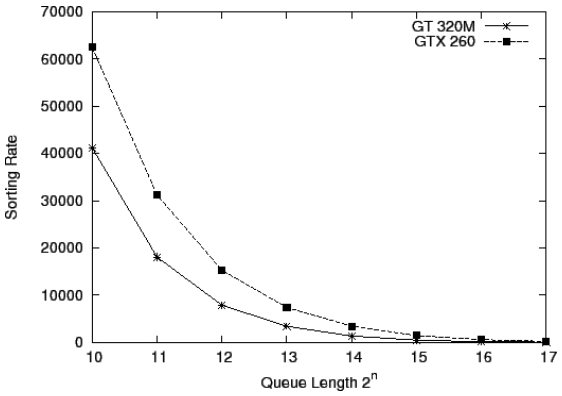
(a) Odd-Even Sort



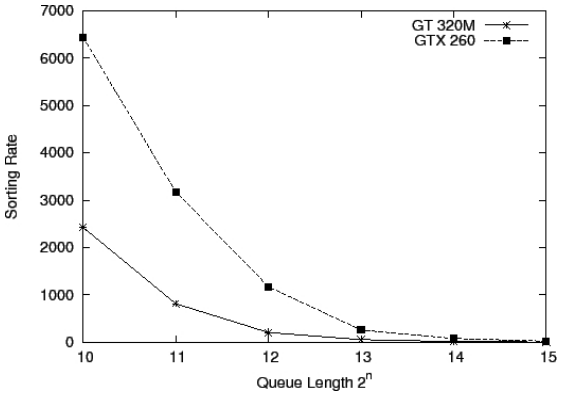
(b) Rank Sort

Figure 3: Sorting Time

3. Related Work



(a) Odd-Even Sort



(b) Rank Sort

Figure 4: Sorting Rate

Sorting is one of the most extensively studied algorithms since well over three decades. Likewise, there is abundant literature on the topic. Since it is not possible to mention all previous sorting algorithms in this paper, we are presenting an occurrence of parallelism in literature only with respect to GPU's in this section. An overview of parallel sorting algorithms is given in [6].

An implementation of the quick-sort algorithm for GPU's using CUDA has been considered in [7] where their results suggest quick-sort as an efficient alternative to radix and bitonic sort for GPU's for larger sequences. For smaller sequences, they suggest bitonic sort. Their quick-sort algorithm uses a divide-and-conquer method for sorting. This involves dividing the sequence recursively into sub-sequences such that all elements lower than a pivot value are positioned to the left and those that are greater are moved to the right. For each recursive call, a new pivot value has to be selected. On the GPU, [7] have proposed two steps; the first being formation of sub-sequences, and the second being allotment of each sub-sequence to a thread for purpose of sorting. The overall complexity of their GPU-quick sort technique is O(n log(n)), with a worst case of $O(n^2)$.

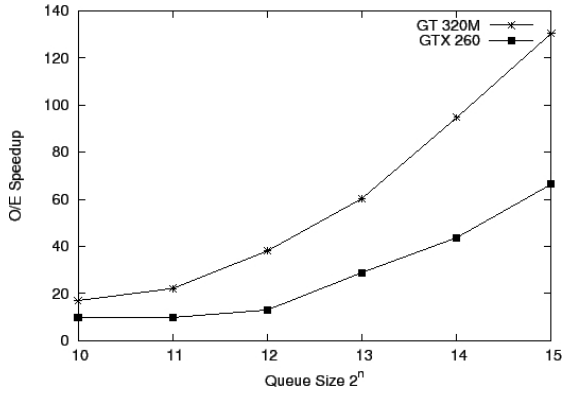


Figure 5: Odd-Even Sort Achieved Speed-up over Rank-Sort

An implementation of radix-sort for GPU's has been provided in [14]. Their radix sort begins by dividing the full sequence *n* into n/p thread blocks where *p* is the available threads. Each sequence is then locally sorted using radix sorting conventions in on-chip shared memory which effectively reduces number of scatters to global memory and maximizes the coherency of scatters. Although scattered I/O is efficient for reducing the amount of system call overhead because a single procedure call is used to write data from multiple buffers to a single data stream, it is not supported in all GPU devices and thus all writes are sequential [15]. On newer cards, including the NVIDIA G80 series and AMD R600 series this is no longer a problem. Their technique achieves complexity of $O(t^{1/2})$ for *t* threads handling 2^b buckets.

Merge soft is also covered in [14] along with radix-sort. Their merge sort follows the same divide-and-conquer approach where the complete sequence is divided into *p* equally sized tiles. All tiles are sorted in parallel with *p* thread blocks using odd-even sort, and then merged together using merge-sort conventions on a tree of log p depth. Their technique is suitable for external sorting, where a processor only has access to a small memory address space. Moreover, since each level of tree are merged together as higher levels are sought, the degree of parallelism is reduced. As a result, this technique may not exploit fully the massively parallel architectures of GPU's.

An adaptive bitonic-sorting algorithm has been proposed in [9]. Their technique sorts *n* values utilizing *p* stream processing units achieving optimum complexity of O((n log n)/p). Bitonic sorting has also been implemented in [13] using the Imagine stream processor.

Sorting queues for traffic simulations have been covered in [16]. Their approach is to study behavior of large groups of transport agents.

4. Performance Analysis

Different C data structures and built-in routines are usually used by programmers for implementing sorting algorithms. In the OpenCL framework however, barring a few supported math functions, most of these are absent. Hence they have to be implemented manually by developers. Moreover, since memory can not be dynamically allocated in kernels, all memory has to be allocated beforehand.

We now examine performance of our sorting algorithms. All performance tests are based on increasing queue sizes where each queue size is a value greater than equal to 2^{10}. The data type chosen for queue sorting are floats. Random numbers for the queue population are generated using uniform and/or exponential distributions. For uniform, the range of values is taken from range 1 and 2^n. The data generated for the queue is performed locally on the CPU, whereas the actual full sorting implementation is done in entirety on the GPU.

The GPU devices used for testing simulation are the NVIDIA GeForce GTX 260 and NVIDIA GeForce GT 320M. The GT320M is designed for laptop systems, consuming low power and has less cores and 1GB device global memory. The GTX260 is a high-end graphics card with 216 cores and 895 MB device global memory. Our results are also compared 2.27 GHz Intel Core I5 CPU with 4GB Ram.

Figure-3 reports the achieved sorting time for different devices. The sorting time is recorded for the actual period spent by the kernel in sorting the queue and does not take into account memory copy and other contention times. On the GTX 260, rank-sort has recorded maximum sorting time for queues size 2^{15} as 47.2s. Whereas equivalent sized queue using odd-even sort has recorded time of 0.7s. On the GT 320M, for queue size 2^{15}, sorting time for rank-sort recorded is 283s or 4m:43s. The same criteria for odd-even sort on the GT320M are recorded as 2.17s. From our results, we see average speed-up of 2.73 for odd-even sort if GTX260 is used. Whereas an average speed-up of 4.89 is achieved for rank-sort for the GTX260. This show that both odd-even and rank-sort will achieve considerable speed-up if number of on-device cores increases.

Figure-5 shows odd-even speed-up over rank-sort algorithm on the GT 320M of 17x for queue size 2^{10} and 130x for queue size 2^{15}. Whereas on the GTX 260 the speed-up of odd-even is 9.7x for queue size 2^{10} and 66x for queue size 2^{15}. The reduced speed-up on the GTX260 even though it has 9x more cores than the GT320 suggests that odd-even sort may have reduced performance edge over rank-sort as the degree of parallelism increases.

Figure-4 depicts the rate of sorting, which is determined as the ratio of queue length and sorting time. For smaller queue sizes, rank-sort has a rounded rate of 2400 elements on the GT320M and 6400 elements on the GTX260. The odd-even sort in contrast shows a rounded rate of 41,000 elements on the GT320M and 62,000 on the GTX260. This gives a rate efficiency in favor of odd-even over rank of 5.8x on the GT320M and 10x on the GTX260. However, in both cases we can observe that the sorting rate approaches 0 as the size of queue increases. Of these, rank-sort converges more quickly than odd-even sort. This suggests that both odd-even and rank-sort do not scale well for large queue sizes.

5. Conclusion

We have presented an analysis of parallel odd-even sort and rank-sort algorithms for GPU's in this paper. Our results show that for small queue sizes, odd-even sort shows a speed-up of up to 5x as compared to rank-sort. Our results also verify that both techniques do not scale well for large queue sizes. And therefore possibility of other sorting algorithms for larger queue sizes, need to be explored further. Our techniques assume that the sequence to be sorted is contiguous

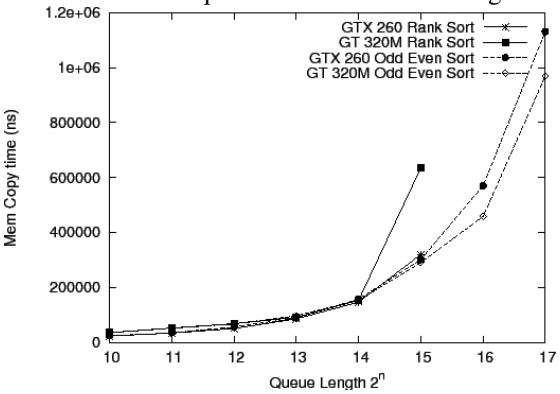


Figure 6: **Memory Copy Times**

arrays of records and therefore the implementation and techniques will not scale well for non-contiguous sequences or multi-dimensional data.

References

3. Ati Firestream DAAMIT GPGPU framework. Website: www.amd.com.
4. Cuda NVIDIA GPGPU framework. Website: www.nvidia.com.
5. NVIDIA GTX 260 specification. Website: www.nvidia.com.

6. NVIDIA Tesla C1060 board specification. Website: www.nvidia.com.
7. OpenCL specification. Website: www.khronos.org/opencl
8. S.G. Akl. *Parallel Sorting Algorithms*. Academic Press, 1985
9. Daniel Cederman and Philippas Tsigas. A practical quicksort algorithm for graphics processors. Technical report, Chalmers University of Technology and Gateborg University, 2008.
10. R. M. DSouza, M. Lysenko, and K. Rahmani. Sugarscape on steroids: simulating over a million agents at interactive rates. In *Agent 2007 conference*, 2007.
11. Alexander Greβ and Gabriel Zachmann. GPU-AbiSort: Optimal parallel sorting on stream architectures.
12. P. Harish and P.J. Narayanan. Accelerating large graph algorithms on the gpu using CUDA. In R. Badrinath, editor, *Lecture Notes in Computer Science*, volume 4873, pages 197-208. Springer, 2007.
13. B. He, K. Yang, R. Fang, M. Lu, N. Govindaraju, Q. Luo, and P. Sander. Relational joins on graphics processors. In *ACM SIGMOD international conference on Management of data*, pages 511-524, New York, NY, USA, 2008. ACM.
14. K. Hegeman, N.A. Carr, and G.S.P. Miller. Particle-based fluid simulation on the gpu. In J. Dongarra, editor, *International Conference on Computational Science*, volume 3994, pages 228-235, 2006.
15. T. J. Purcell, C. Donner, M. Cammarano, H.W. Jensen, and P. Hanrahan. Photon mapping on programmable graphics hardware. In *Annual ACM SIGGRAPH/Eurographics conference on Graphics Hardware*, pages 41-50, 2003.
16. Nadathur Satish, Mark Harris, and Michael Garland. Designing efficient sorting algorithms for many core GPUs. In *23rd IEEE International Parallel and Distributed Processing Symposium*, May 2009.
17. Shubhabrata Sengupta, Mark Harris, Yao zhang, and John D. Owens. Scan primitives for gpu computing. In *Graphics Hardware*, 2007.
18. David Strippgen and Kai Nagel. Using common graphics hardwaer for multi-agent traffic simulation with cuda. In *SIMUTools*, 2009.

Innovation of the Software Product PROMAN W®

Daniel Petrik
MMS Softec ltd.
Trnava, Slovakia
City, Country
e-mail: petrik@mms-softec.sk

Oliver Moravcik
Institute of Applied Informatics and Automation
Slovak University of Technology
Bratislava, Slovakia
e-mail: oliver.moravcik@stuba.sk

Abstract[1]

Product Development of PROMAN W began in 1999 and the product was until last year, updated regularly. During its development it was used in client-server architecture. The client was designed as a Windows application, communicating with clients SQL database server (Oracle, MS SQL). The application was written in a developing environment – the Delphi
language Objective was Pascal.
Due to the further development (implementation of customer requirements) and maintenance becoming more challenging, a decision to re-engineering the product was reached.
Set requirements of the new technology:

- Strict usage of objects, no SQL commands in the application.
- Separation of user interface (UI) from the "code" application.
- Building a user interface (UI) based on visual object (controls), easy customisation using templates and styles.
- Possibility of automatic testing.

Set requirements for use of a modern and comfortable developing environment with the possibility to generate the largest part of the source code according to the UML model. Reverse engineering of source code to the UML model (the actualisation of the UML model from the source code).
Due to customer needs, it was decided that the prime platform (operating system) would remain as MS Windows, the following was done in order to meet the following requirements:

- Object-relational database mapper (OR Mapper) - provides persisting object to relational database, management of the database structure. Database servers are supported- Oracle, MSSQL and Firebird.
- Composite Application Library (formerly Prism) – allows the distribution of applications on its own modules, allows a loose coupling between modules, delivers a pattern Model-View-Presentation Model, separation of presentation and business logic of applications from UI logic.
- WPF (Windows Presentation Foundation) – allows the building of an object-UI, an easy connection of UI to the presentation model with automatic direction, actualization (Data Binding), defining and use of templates and styles to change appearance of UI elements etc.
- Due to the application of the pattern MVP, it is possible to use unit tests to test a presentation model and model.

To create a UML model and to generate source code "Enterprise architect" was used. It is also used to update (Reverse engineering) the UML model (class diagrams) according to source code (C #).
For development and testing, a development environment (MS Visual Studio 2008) is used.

1. Introduction

Product Development of PROMAN W began in 1999 and the product was until last year, updated regularly. During its development it was used in client-server architecture. The client was designed as a Windows application, communicating with clients SQL database server (Oracle, MS SQL). The application was written in a developing environment – the Delphi
language Objective was Pascal.
Due to the further development (implementation of customer requirements) and maintenance becoming more challenging, a decision to re-engineering the product was reached.
Set requirements of the new technology:

- Strict usage of objects, no SQL commands in the application.

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

- Separation of user interface (UI) from the "code" application.
- Building a user interface (UI) based on visual object (controls), easy customisation using templates and styles.
- Possibility of automatic testing.

Set requirements for use of a modern and comfortable developing environment with the possibility to generate the largest part of the source code according to the UML model. Reverse engineering of source code to the UML model (the actualisation of the UML model from the source code).
Due to customer needs, it was decided that the prime platform (operating system) would remain as MS Windows, the following was done in order to meet the following requirements:

- Object-relational database mapper (OR Mapper) - provides persisting object to relational database, management of the database structure. Database servers are supported- Oracle, MSSQL and Firebird.
- Composite Application Library (formerly Prism) – allows the distribution of applications on its own modules, allows a loose coupling between modules, delivers a pattern Model-View-Presentation Model, separation of presentation and business logic of applications from UI logic.
- WPF (Windows Presentation Foundation) – allows the building of an object-UI, an easy connection of UI to the presentation model with automatic direction, actualization (Data Binding), defining and use of templates and styles to change appearance of UI elements etc.
- Due to the application of the pattern MVP, it is possible to use unit tests to test a presentation model and model.

To create a UML model and to generate source code "Enterprise architect" was used. It is also used to update (Reverse engineering) the UML model (class diagrams) according to source code (C #).
For development and testing, a development environment (MS Visual Studio 2008) is used.

Proman W® is an application with a main focus on project and personal costs management, including budget planning and balancing. It is an open application, equipped to reach a number of interfaces. These interfaces allow one to read and prepare data for cooperating applications i.e. the following accounting systems (MACH, SAP, WinLine) or personal accounting systems (Best, PAISY). Proman W was designed as a client-sever application with support for MS SQL and for the Oracle SQL server and was developed in Delphi (first RAD tool using object Pascal). The first version of Proman W was released in the year 1999 and the product still exists to this day (year 2010).
In the year 2008 we decide to re-engineering the Proman W application in order to allow further evolution of the product (customers-users do not like revolution!). We did research for actual technologies to define the base technological requirements, while user requirements (user experiences) were already known.
As a result we got the following technological requirements:

- Use of data objects for data access, no direct access to a relational database. The goal is to shield the developer from SQL commands and its' database implementation dialect. (A side effect is an enlarging number of supported database systems).
- Use of the UML model for application modelling and code generation, actualisation of the UML model from code (reverse engineering).
- Application client support for Windows platform (32 and 64 bit).
- Use of object "like" composition of user interfaces (support for styling, templating).
- Open path for later support of web-like application.
- Maximization in the reuse of application parts (modules).
- Automation of a test process (where possible and effective).
- Use of a user friendly development environment with support and an ensured future.

2. Fulfilments of technological requirements

2.1. Use of a user friendly development environment with support and an ensured future

Having developer experiences with IDE like Delphi, MS Visual Studio and Centura (Gupta) we now know what we have to expect from the modern development environment. We used the supported platform MS Windows, for the year 1999, Borland Delphi in 2007 and unfortunately in the year 2008 MS Visual Studio 2008 due to a much more comfortable development environment, especially because of the code editor's intelligence and debugging with the visualisation of object data. Because we already have experience with C# (web-application based on aspx-pages), we decide to use MS Visual Studio 2008 for further application development.

We also keep in mind what happened with Kilix (which we were using to support a cross platform application – Win32/Linux) for a few years.
Actually we have updated the developer environment to MS Visual Studio 2010. The main reason for this was to improve the support for team's source management, such as work items, bugs, builds, and source code versioning.
The application using C# uses framework .NET, so they are built as a managed application. It is promising that the developer of the managed application does not need to carry out memory freeing (this is done automatically by the garbage collector), but the reality is that the developer must prepare a condition for it (more info in Modern Trends in the Development of Software Applications). To check memory management we use the tool YourKit Profiler, to profile memory use and possible memory leaks together with its' reasons.

2.2. Use data objects for work data access

The simplest realisation of this requirement is to use an object-relational database mapper. Actually there are many products and a decision without prototype testing is difficult. The first relevant information we have found was in Diploma thesis "Objektově-relační mapování pro platformu .NET" (Object- relational mapping for platform .NET) from I. Stanek (May 2008, CVUT Praha). The next requirement – uses the UML model to generate data objects and a shortened list of products, while some of them can generate data objects only from database objects, but not the opposite. An additional test also shows that proper configuring is not a simple process (NHibernate). Our final decision was to use the product "eXpress Persistent Objects" (XPO) from Developer Express Inc, this was in the year 2008. This product covers ours requirements and its license rules and price was acceptable.
When using XPO we can almost forget about a relational database while this OR mapper creates and modifies database objects. Data objects are automatically instantiated from data classes upon an application's request. Data classes are defined by the UML class diagram and automatically created by tool Enterprise Architect. This saves much of the developer's time and allows the UML Class model to stay synchronised with the source code.
During development we found that use of the XPO OR-Mapper has also got some disadvantages. One is decline of performance when deleting the root object of an object's tree. This happens because all child objects should be first read into memory and then marked as deleted. This is a general problem of OR-Mappers. Also unsupported calling of stored function/procedures can create a problem. Stored procedures are sometimes used to copy bulk data between tables (creating data snapshoot). There is a solution – this is by calling the stored procedure indirectly as an action triggered by the database "control" object property change.

2.3. Use of object "like" composition of user interfaces

While our application is dedicated to the Windows platform we have to find a solution, where visual components would be defined and behave like objects, with a possibility to simple modify the visual look of a component whilst retaining its functionality, e.g. something like a "lookless" visual control. If the control will support the modification of the visual look via a template or style, it would also be a big plus (something like css know as cascade styling in the world of web pages).
The matter of course is to support different display resolutions, especially with high DPI (most of the actual application in this case shows either very little fonts or a corrupt window layout).
Due to the above mentioned, we decide to use Windows Presentation Foundation (WPF) as a foundation for our UI. A very strong feature of WPF is data binding. This allows a response of WPF user controls, not only on user gesture, but also on a change of bound data. The presenter does not have to care about the refreshment of a visual component, it just sends a notification of change, usually automatically raised when there is a presenter property change. This notification can work in both directions, so a presenter can also be notified that a user modified some data in the visual component. WPF has also modified layout system, where each visual control knows how much space it requires to successfully render. This allows the dynamically updating of window layout. Due to it almost being possible to update each visual component i.e., to apply a transformation, zooming of the window or the whole, the application is very simple.
Another plus is the number of existing third party components as intelligent data grids with filtering sorting and grouping.
Our application consists of the main window (shell) which contains the main application window with a menu. Menu items are bound to a common command, where command parameters define the presenter name, which should be created and activated. This is done by the main menu presenter. Presenters are registered in their modules together with their main views.
The WPF's point of view as a user control is then shown in the main application window.
Because the view itself is the user control, and it contains other controls and user controls, we can build a view like a puzzle.

3. Maximizing the reuse of application parts (modules)

When defining the application's architecture we decide to implement the idea of a separation from the visual

part of the application and from data and its processing. We study patterns used for this purpose and after developing a few prototype modules, we found out that the pattern Model-View-Presenter (MVP) is most fitting for our situation.
This model is responsible for basic data operations such as: reading, storage, updating and deleting, it also implements business and validation rules. The model contains and centralises all functionality related to data processing.
View – is responsible for the rendering of visual controls, interaction with the user (user input via keyboard, mouse, etc.) and for rendering the data from a presenter. The View can contain some logic relevant for visual components, i.e., change of background colour as a response of visual data and state of change. This is possible because WPF supports triggers, triggered by data change or visual component property change. View can also do some visual data transformations using converters and data templates.
Presenter – administrates all user events forwarded by View, transforms data (data objects) from the model's view and transforms user interaction events to the calling of data action of the model. The Presenter, as we use it, is focused on interaction logic, which was not possible or difficult to realise in View.
This separation allows the reuse of models, e.g., for import/export interfaces. For example the importing of data from an external application must fulfil the same rules as entering data into an application, which is assured by the use of the same model.

4. Open path for the later support of web-like application

The web-like application uses a different application scenario – web client – web server – database server. Use of a MVP pattern allows the sharing of model modules, while modified the presenter and View is realised as an "aspx-page". Of course this solution is not dedicated for a high load application, while the model is not optimised for this situation, but is sufficient in a scenario where there the web browser is used as an application by clients.

5. Automation of the test process (where ever possible and effective)

The test process as we understand it can split into:

- Testing of methods – we use unit tests to test methods, writing a full test is very complex and time consuming (to check all possible cases – at least for boundary parameters).
- Testing of model methods – in most cases the result of this test is the modification of data in a database. This complicates testing, while we need to ensure an identical state at the beginning of the test. This is solved using an ordered test or (and) using a data provider in memory when XPO data objects are saved in a memory database.
- Test of the presenter – this is similar as the test of model methods, but there is an extended testing of the property value as a result of commands (called from the test).
- Test of view – we have not successfully covered this part, writing test scripts is a very time consuming operation, it seems that the best method is to capture user interaction, then to repeat it by the use of a test engine and then to check the visual state of the elements and the presenter properties.
- Integration test – we do this manually.

6. Conclusion

Our experiences have shown that the correct choice of the technology becomes time a consuming operation. It also shows that without prototype testing it is hard to correctly evaluate any new technology.
We try to mention some modern technologies used in the new Proman W application together with the advantages and problems when implemented into a real application.

Approach to prognosticating the Evalution of Program System

Evgeniy Reshetnikov, PhD Assoc. Prof.,
Head of Department of Organization of Computational Processes and Control Systems
STU Izhevsk (Votkinsk Branch)
Russia, Votkinsk
e-mail: beevasya@gmail.com

Abstract

The article provides a brief description of assessment costs for the development of software systems, describes some of the shortcomings and suggest alternative methods of assessment, based on measurement of Architectural-program complexity of the software

Introduction

At present, we can observe rapid growth of software industry. This is due to the emergence of many organizations involved in software development. However, many organizations are faced with the problem of prediction and assessment of complexity of design. The accuracy assessment will affect the competitiveness of the organization of the developer, because high accuracy will allow more efficient approach to the formation of the final cost of the software product.
Currently, there are a number of methods for assessing the value of your software: COnstructive COst MOdel (COCOMO), SEER for Software (SEER-SEM), Constructive Systems Engineering Cost Model (COSYSMO), etc. One of the most popular techniques is COCOMO. The method allows COCOMO estimate the complexity and time of software development. It was first published by Barry Boehm [1] in 1981 as a result of the analysis of 63 projects of the company «TRW Aerospace». In 1997 the technique was refined and became known as COCOMO II [2]. Calibration of the parameters was performed on 161 development project. The model uses the regression equation with parameters determined on the basis of sectoral data and the characteristics of a specific project. The presented method is based on such a project as an indicator of the size of the product in thousands of lines of source code (KSLOC, Kilo Source Lines Of Code). The size of the software product is determined by the expert, for example using the technique PERT [3].

Publishing

The main problem of all available methods is the procedure for determining the objective of the program in lines of source code. To ensure accuracy in determining the size of a certain number of experts required sophisticated skills. For some organizations it is difficult to fulfill such requirements, which consequently will lead to less accurate estimates. We can therefore say that the existing methods are most effective in large companies employing more than 100 people. At the same time in the software industry there is a sufficiently large number of companies with employees numbering less than 50 people. For such organizations existing methods of valuation software are not effective. It is therefore proposed the development of alternative methods of estimating software, which allowed to exclude the described drawback.
As a basis for the development methodology is proposed to use some indicators of Architecture and software complexity, which is a quantitative characteristic of the software. Indicator architectural complexity of the program will be determined as the sum of the complexities of all the elements comprising the software. Figure 1 provides a general description of the methodology.

$$C = \sum_{i=0}^{n} E_i \ ,$$

where *C* - an indicator of the complexity of software architecture, *Ci* - the complexity of the i-th element.

$$C = f(K_1, K_2, K_3, ..., K_{n-1}, K_n) \ ,$$

where *K1, K2, K3, Kn-1, Kn* factors affecting the complexity of the architectural program element.
It is expected to formalize the set of basic architectural elements of the program. Each element will be characterized by its own set of coefficients. Information on the architectural program elements will be collected from some of the design documentation for software as design documentation can be used, for example, UML-diagram.
Immediately importance in development costs will be determined on the basis of a linear dependence.

$$E = a*C + b,$$

where E - meaning cost, C - Architecture and software complexity of the software, a and b - regression coefficients.

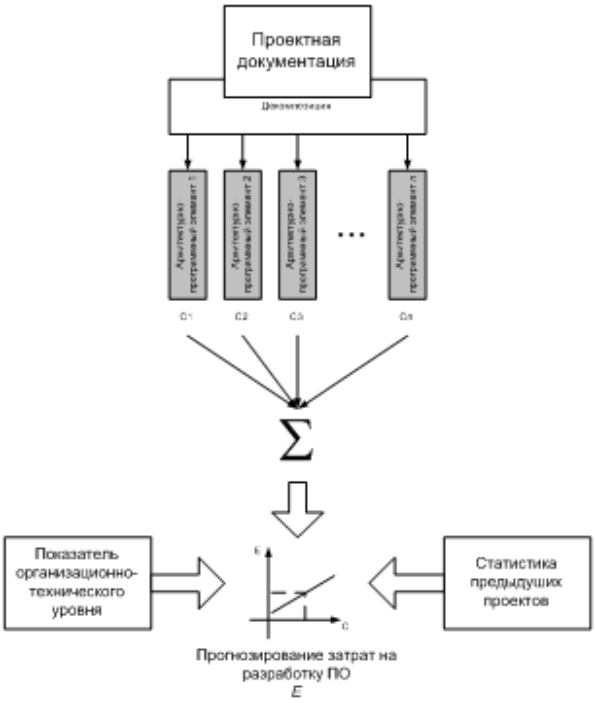


Figure 1. General description of the methodology

Regression coefficients are some indicators of organizational and technical level of the company, which is assessing the cost of software development. The coefficients can be determined in two ways: the statistical (based on the experience of previous developments) and analytical (based on the performance of the company, for example, the availability of certain tools, the experience of developers and others). As part of the described work it is planned to develop methods for determining the organizational and technical level of the company.
Thus, for the implementation of the proposed method to solve two main tasks:
1. formalize the many types of architectural and software elements, and for each type of element to form a list of coefficients;
2. develop a method for analytically determining the organizational and technical level of the company to obtain the regression coefficients.

Conclusion

As a result of scientific work the method of estimating the costs of software development, which will be effective for large organizations and small software development companies, will be presented. As a practical result will be the creation of an automated system for calculating the cost of software development.

References

1. Barry Boehm. Software engineering economics. Englewood Cliffs, NJ:Prentice-Hall, 1981. ISBN 0-13-822122-7
2. Barry Boehm, Chris Abts, A. Winsor Brown, Sunita Chulani, Bradford K. Clark, Ellis Horowitz, Ray Madachy, Donald J. Reifer, and Bert SteeceSoftware cost estimation with COCOMO II (with CD-ROM). Englewood Cliffs, NJ:Prentice-Hall, 2000. ISBN 0-13-026692-2.
3. [Electronic resource] // PERT chart - URL: http://www.netmba.com/operations/project/pert/ (25.03.2010) Free Access

“Joy-of-Use” and “Joy-in-Error”

Kerstin Palatini
Anhalt University of Applied Sciences
Department of Computer Science, Köthen
e-mail: kerstin.palatini@inf.hs-anhalt.de

“Everything that can be counted does not necessarily count; everything that counts, cannot necessarily be counted.”
(Albert Einstein)

Abstract

The paper deals with “Joy-of-Use” as a concept of the *Human-Computer-Interaction (HCI)* and its chance to develop a new error culture, if we accept new thinking about and new dealing with errors.

Keywords: “Joy-of-Use”, Usability, User Experience, Error, Rules and Norms, Design, Computer Science, Interdisciplinarity, Culture

1. Introduction

“Joy-of-Use” is close to the concept of usability and user experience in HCI. It was established in the years when the computer became a daily used tool and medium, especially by using it at work but also more and more at home, e.g. by using the wide range of internet services.
HCI deals with new questions and problems of interaction and how to make using computer based artefacts and techniques easier and more intuitive.
Usability (as a combination of the words use or usable and ability) together with the user experience as a special field of HCI have enjoyed acceptance and an increasing significance in scientific research and economy since the 1980ies.
Since that time experts have worked in this field and have developed rules and norms to make usability possible and easy to achieve.
Human and emotional factors are a little problem of achieving safe usability and user experience.
Errors in scientific fields like in computer science are suppressed by rules and norms.
But is this really the best way of dealing with “unexpected interactions” [1].

2. Usability, Joy-of-Use and User Experience

2.1. Usability

Usability describes whether the product, the software or the website is user friendly. According to the norm DIN EN ISO 9241 something is user friendly, if it can be used efficiently, takes less time and resources to accomplish a particular task effectively and to achieve user satisfaction. Usability comprises attributes that assess how easily artefacts are to be used. The word "usability" also refers to methods for improving the Ease-of-Use. Usability consultant Jakob Nielsen [2] and computer science professor Ben Shneiderman [3] wrote about a framework of system acceptability, where usability is a part of "usefulness" and it is composed of Learnability, Efficiency, Memorability, Satisfaction and Error Tolerance.

2.2. “Joy-of-Use”, User Experience is more than Usability, but they are nothing without usability

„Joy-of-Use“ is a concept based on HCI and usability but particular attention is layed on the experienced and emotional quality of dealing with artefacts. Different theories have been developed almost simultaneously [4]. Today the main and broad concept of these different theories is reflected by the User Experience (UX). So we can say: “Joy-of-Use” is part of user experience. User experience is more than usability, because user satisfaction is more than just functional aspects and the absence of stress.
But “Joy-of-Use” without usability doesn’t work.
Prof. Irrgang (TU Dresden, Technical philosopher) says that to buy something - like a Hi-Fi system - just for fun and aesthetics but not to use it is nearly subversive. [5]

3. “Designing Joy”

3.1. Usability Engineering- by far more rules and norms – “Engineering Joy”?

The process to “construct” Usability is the so-called “Usability Engineering Process”. ISO 13407 describes this process as a Human Centred Design Process for Interactive Systems.
The norm describes the necessity of
- Multidisciplinarity

- Implementation of design in the process
and the
- User Centred Orientation from the very beginning of the process.
Further we can read in ISO 9241-10:
suggestions for about (at least) 10 Ergonomic Dialog-Requirements, e.g.
- Suitability for the task
- Self-descriptiveness
- Controllability
- Conformity with user expectations
and one of them is:
- Error Tolerance (we come back to it later)

These and a lot of other norms and guidelines [6] were worked out to achieve usability in an engineering process and as "operating instructions".
Knowing that and knowing on the other hand that we must work on meeting user satisfaction, what means: develop for emotional needs, for joy – we can really doubt that only rules and norms will help us to "install" or "engineer Joy".

3.2. Joy as an Emotional Quality

We know the three main requirements from the definition of usability: effectiveness, efficiency and satisfaction.
So we have an example of functional and so-called "non-functional" requirements (Author means that it is better to say subjective and emotional requirements than "non-functional", because it has been proved that the functionality of emotions in our days is very important.)
Like other "non-functional" requirements, usability cannot be directly measured but must be quantified by means of indirect measures or attributes such as, for example, the number of reported problems with the Easy-of-Use of a system [7].
So emotions like joy cannot be directly measured. It is possible to compare them in such a way like before and after, yesterday and today; you can measure physical processes like the forced brain activity and visualize it in an attractive way but you can´t measure joy reliable.
And secondly: if we think about the factors influencing our personal emotional state (e.g. age, gender, own experiences, culture, knowledge, weather ...), so we can really doubt the reliability of our assessment.
And we imagine, that we cannot "design joy" so easily.

3.3. Interdisciplinarity in this process: for example between Designers and Computer Scientists

One of the claims of ISO 13407 is multidisciplinarity.
According to the complexity of interactive computer based systems several experts in different fields have to work together; for instance, industrial psychologists, ergonomists, computer scientists, and designers.
ISO 13407 requires in particular above all (in any case) the cooperation between designers and computer scientists.

So let's just have a look on some differences between these two disciplines and their way of thinking and solving problems.

Computer scientists have a strong mathematical base and education; they learn thinking in schemes, in algorithms. The solutions are based on rules and norms.

On the other side the designers learn to work with creative methods to develop unusual things, to develop new or innovative systems that catch our attention and make every day life more comfortable.
They also have rules but they know: if you want to structure really new things often you have to break down these rules.

And if we also know that different ways of thinking lead us to different actions we get a first answer a) to problems existing between the two disciplines that have to develop usability and user experience simultaneously and b) we can imagine our different ways of thinking (and dealing) with errors.

4. Dealing with Errors

Errors may have a different meaning.

4.1. In Computer Science

In computer science an error is a failure in hard- or software or in interface design and it has to be debugged as soon as possible. [8]

If we imagine or think of examples of errors in computer systems, so we think of catastrophic scenarios especially in powerful systems (in a car, in an airplane, in a nuclear-power-plant). Today it is also an economic task for business companies to make their systems "fail-safe".
Anyway errors are not desirable for computer scientists and their customers, and we cannot notice "Joy-in-error" in this field (exactly in contrast to such "Joy-in-error" for "computer specialists", who develop criminal energy to crack systems or work with enthusiasm to create new errors, new virus attacks).
There are theories of error-management in computer science (e.g. "Failure Mode and Effects Analysis (FMEA)"), but they are not exhaustive. [8]
So we can find some functions to control errors with so-called error-routines in context with the call for "error-tolerance" of a system.
An example for an error-handling-function is the "On-Error" instruction in the programming language Visual Basic [9].

Sub fehler2()

```
' ignore error

On Error Resume Next

x = 5 / 0


' Error Handling off!

On Error Goto 0

x = 1 / 0

End Sub
```

With this method you can "catch or ignore errors", "jump over errors" or "cut-off the error-routine".
It is a so-called "structured exception" to uncontrollable data or state in a program flow and non-calculable behaviour of the user.
But is it not an illusion to have a tool against the "non-calculable behaviour of the user"?
It give developers (and their customers) the imagination of absolute controllability of a system (inclusive the behaviour of the user).

It is not astonishing to hear in technical disciplines – like in computer science – the words: "Only that what you can measure you are able to control." and this sentence could be (dangerously) continued as follows: "And that what you can't measure you cannot improve."
But is it really the genuine truth?

1.) On one hand we know that the same (measured) facts can have a very different meaning depending on appraisal and contextual supply linked with a subjective (often goal- or target-oriented) interpretation [10] and on the other hand
2.) we all know, what things can mean to us in a positive (or negative) way without knowing all the data, without knowing all the numerical and objective facts about and behind them, without knowing the not-measurable quality; for instance of beauty, love or joy.
A wise man expressed in a much better way:
"Everything that can be counted does not necessarily count; everything that counts cannot necessarily be counted." (Albert Einstein)

4.2. In Design

Dealing with errors in Design means to find out errors and mistakes in product properties and product handling and to manage them for a better and a more likeable use.
Problems are challenges to create new solutions.
Regarding errors and mistakes one of the challenges for designers is to find and discover chances to make things better, to enrich our everyday life.

In this case designers have the right to doubt. At first we notice the problems and errors, then we have to ask questions and these questions bring about chances, requirements and changes [1].
So an error has definitely positive dimensions for this discipline.
Often in the design context new ideas are required and *innovations* are challenged.
But even innovations have a high potential of risks and errors.
Keith Pavitt [11] defines in his characteristics of technological innovations (inter alia) that they involve continuous and intensive collaboration and interaction between functionally and professionally specialized groups and also that they have profoundly uncertain activities that remain. It also means that only one of ten turns out to be commercial success.
Uncertain activities also mean to deal with risks and errors.
We see: innovations cannot come into existence without making mistakes.

In design we need people who know and love such challenges. This is not as usual as it seams to be, because quite often work is only aimed at lower costs and higher profit.

4.3. Perspectives of "Joy-of-Use" and "Joy-in-Error"

To create "Joy-of-Use" the disciplines have to collaborate. This work is determined by different ways of thinking and dealing with errors.
We can only work successfully, if we know the different paradigms and accept the different approaches and concepts and combine them with a new understanding.
But we must also involve new paradigms and perspectives when handling errors.
We also have to notice and accept that there are several categories of errors.
Another step is to accept that we have to live with errors.
We as human beings are incomplete. The philosopher Karl Popper said, „It is nothing else but the unvarnished truth that we make mistakes. " [12]
We have to learn to live with errors that we can handle in a constructive way.
It requires thinking and talking about errors and analyzing the problems permanently and then looking for something better. "Trial and error" is a hard practice (but the main impetus of our society). It needs courage, knowledge, power and, of course, the awareness that there ever will be risks.

First of all we must change our opinion about errors.
Change the paradigm of errors under the keyword "new error attitude".

Conclusion

The concept of "Joy-of-Use" and Usability accepts that there are functional (quantitative) and "non-functional" (qualitative) aspects important to develop and use technical (computer based) artefacts.
For Usability Engineering complexity and innovation mean an interdisciplinary teamwork, at least by designers and computer scientists.
They have very different ways of thinking and dealing with rules and norms and therefore of their handling of errors too.
A new thinking and talking about errors can help to establish a new error-attitude.

To develop useful and joyful products in the usability engineering process could present an adequate platform or interface for practical testing and error would be based on a new understanding of errors.

The user also should learn: if he wishes to participate in new techniques and systems he has to collaborate.
He also should abandon perfection.

But all these challenges need broader acceptance and a better understanding of error in business, economy, policy and society.
It is hard enough to demand rethinking in our society characterized by an anti-failure culture.

Technical-philosophical aspects about the development and dealing with products and systems can help us to analyze the errors and the inherent possibilities of change. In this context it is important to discuss ethical aspects and the meaning of trust and communication.

Let's finish with a quotation of Wolf Lotter [1]:
„To err is human.
Insanity too.
Who nothing attempts who will not become wise.
Only idiots believe to be perfect."

References

1. Lotter, W. „Fehlanzeige". In: *brandeins* 08/ 2007 Wirtschaftsmagazin, Medien AG, Hamburg 2007.
2. Nielsen, J. "Designing Web Usability: The Practice of Simplicity". New Riders 1998.
3. Shneiderman, B. "Designing the User Interface: Strategies for Effective Human-Computer Interaction". Addison-Wesley, Massachusetts 1992.
4. Reeps, I.E.: „Joy-of-Use eine neue Qualität für interaktive Produkte". Universität Konstanz 2004.
5. Irrgang, B. „Philosophie der Technik". Darmstadt 2008, p. 317.
6. see: http://www.handbuch-usability.de/normen-und-standards.html;http://www.einfach-fuer-alle.de/blog/id/975/ (15.6.2010).
7. Hassenzahl, M., Beu A., Burmester, M. (2001). „Engineering Joy". IEEE Software [1&2], pp. 70-76.
8. Richter, V. „Joy on Error". 3rd International Workshop *„Innovation in Information Technologies*", Dresden 2010.
9. http://msdn.microsoft.com/de-de/library/5hsw66as(VS.80).aspx (18.6.2010).
10. Steinert, St. "Visualisierungstechnologie und "Mixed Hermeneutics". In: Irrgang, B. (Hrsg.) *"Dresdner Studien zur Philosophie der Technologie"*. Bd.3, Peter Lang Verlag der Wissenschaften, Frankfurt /M. 2010, pp. 91 – 108.
11. Pavitt, K. "Technology, Management and Systems of Innovation". Edward Elgar Publishings, Massachusettes 1999, p. 63.
12. Popper, K. "Die offene Gesellschaft und ihre Feinde". Francke Verlag, Bern 1958, II. Bd., p.466.

An Overview of Methods for 3D Model Reconstruction from 2D Orthographic Views

Michal Elias Michal Kebisek

Institute of Applied Informatics, Automation and Mathematics
Slovak University of Technology in Bratislava
Trnava, Slovak Republic
e-mail: michal.elias@stuba.sk e-mail: michal.kebisek@stuba.sk

Abstract[1]

In this article, an overview and analysis of known methods for 3D model reconstruction from 2D orthographic views is presented. These methods deal with automatic processing of technical drawings of parts in vector format. The subject is a part of a complex problem of transformation of paper drawings to digital three-dimensional model. It is based on the idea that most of the design work doesn't lies in designing new components, but in adapting, modifying and refining existing ones. This issue is not new, but it is still current, as demonstrated by many articles that deal with it.

1. Introduction

Reconstruction methods developed by many researchers can be classified to two main branches.
Earlier approaches are based on Boundary Representation of 3D objects (B-Rep). They transform the 2D points and lines of orthographic views to 3D points and lines. The lines must be verified and classified to eliminate duplicate or illegal lines and points. A wireframe model is created. Then, from line segments the faces are constructed and finally B-Rep 3D object is build.
Newer approaches are based on adding and removing material and are therefore closer to machining technology. They use Constructive Solid Geometry (CSG) to construct a 3D solid model. In CSG a binary tree is build where nodes represent Boolean operations (addition, subtraction, union) and leaves represent primitive objects. Some methods start with a boundary cuboid and step by step remove overflowing volume. Sophisticated methods are searching for partial projections of primitive elements (features) and build a 3D solid by adding, subtracting and union of base and complex volumes.
There are millions of technical drawings produced previously in paper or digital vector form, but modern CAD/CAM systems works only with 3D models. In this overview we try to find the best solution of 3D model generation from 2D projections and help to increase the usability of old technical drawings in engineering process of new products.

2. Boundary Representation (B-Rep)

The pioneer in the automatic reconstruction of 3D objects from 2D orthogonal projections is Idesawa with his paper [13] from 1973. His method is based on the bottom-up approach of boundary representation – generate 3D vertices and edges, construct wireframe model, eliminate false edges, generate faces, eliminate false faces, assembly 3D solid object. This approach was limited to reconstruct only rectilinear polyhedrons. During following years, other authors improved his method and so extended the set of 3D objects that can be reconstructed. Processing of rounded edges and cylindrical and conical surfaces was added also.
Sakurai and Gossard in article [22] first addressed the possibility of generating curved surface formed by circular arcs. They advise of tangency edges which lie between curved and flat faces and are not visible in orthogonal projections. Vertices are categorized in several types e.g. standard, tangency, and silhouette. This categorization in pre-processing of views significantly simplifies further processing.
Dutta and Srinivas developed a method for reconstructing rounded objects using only two polygonal orthogonal projections [7]. The first step will generate all possible third views, which correspond to the initial two views. Next, these new views are analyzed. Line segments that can be replaced by circular arc are recognized. Finally, the curved solids are generated which correspond to the pair of input projections.
In article [14], the authors Itoh and Suzuki described a method to eliminate pathological vertices, edges and

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

surfaces that come from the creation of surface model. It is based on Boolean evaluation criteria.

You and Yang developed a method for the reconstruction of B-Rep model including flat and curved surfaces [28]. Like everyone else, using a boundary representation, they start with a wireframe model and deal with the removal of pathological edges and surfaces. They don't generate all possible combinations of lines and faces, but only those that make sense to create a manifold.

Masuda and Numao presented an effective method based on cell basis [19]. The projections will create a single wire-frame model of non-manifold, which is the basis for finding the different candidates for the final solid model. Cells that are bounded by various combinations of faces are recognized in the wire-frame model. A test is performed after removal of the cells whether the resulting projections of the model coincides with the input projections. They deal with the errors introduced by man in the projection drawing. Their method can correct and add missing projections of edges and incorrectly used line types.

Shin and Shin in the paper [23] presented a method based on principle first introduced by Sakurai [22]. They reconstruct 3D solid objects using geometric properties and topology of geometric primitives.

Watanabe in article [25] presented a method, working on B-Rep principle by which he can handle inconsistent input. He recognizes redundant line segments, inappropriate line types and the lack of line segments. This method is applicable in the processing of drawings gathered by vectorization of scanned paper drawings.

Liu et al. developed a method [18] for the reconstruction of objects containing the conic surface which axes may be oblique to the axis of coordinate system, and thus are reflected distorted.

Chovan in his thesis [4] dealt with the generation of three-dimensional wire-frame model of an polyhedron. He developed functions for pre-processing of input data such as alignment of projections and the elimination of defective lines. He also addressed the problem of pathological features that were mentioned by all his predecessors, who used the boundary representation of 3D model.

Pavelek developed a stand-alone application that reconstructs an surface model. In his thesis [21], he carefully analyzed the rules for the orthogonal projections and he design and implement the reverse approach. Ambiguous and inexact inputs were also processed.

3. Constructive Solid Geometry (CSG)

Constructive Solid Geometry was first used by Aldefeld [1] ten years later after Idesawa. Since then, the CSG approach was chosen by many researchers because of its speed and simplicity of 3D solids reconstruction and new methods were developed and introduced. These methods are also known as top-down approaches.

CSG method used by You and Yang [27] starts with base boundary block. Feature parts are step by step recognized, then reconstructed by profile sweeping and finally removed, until the goal 3D solid is obtained. They introduce use of auxiliary views and dimensions and geometric tolerances extraction.

Human approaches to engineering drawings interpretation were involved in method described by Geng at al [10]. Using three projections views a 3D cell-box is build. They decompose the 3D object in a set of 3D cells; every cell is generated by temporary use of wireframe algorithms. The reconstruction process starts from bottom and proceed layer by layer to the top of the cell-box. For each layer of cells the best usable operation and direction is used. The final solid model is constructed by union of the sub-objects.

An interesting method is proposed by Cicek and Gulesin [5]. In the first stage, feature parts are constructed by extrusion and revolution operations from inner entities evaluated on three orthographic views. In the second stage, outer profiles are extruded using information from neighbouring views to obtain three basic volumes. Feature parts are then subtracted from the basic volumes and the final solid model is obtained by intersecting of all the three volumes.

Reconstruction of 3D Solids of revolution is proposed by Lee and Han in article [16]. Efficiency of CSG is used to represent the revolution with interacting volumes so more complicated objects like shafts can be reconstructed.

Lui and Ye presented a method based on engineering semantics understanding of technical drawings [17]. From projection views is extracted graphic feature information. Then, engineering semantics is used to recognize primitives, components and the relationship between them. Corresponding algorithms are used to reconstruct every component. Finally, all primitives and components are assembled into a goal solid using Boolean operations and CSG. Every new kind of reconstructed primitive is saved as new pattern to increase efficiency of the method.

Benkovsky dealt with reconstruction of 3D solids of simple revolution parts from one projection view [2].

A boundary cuboid is used by Cayiroglu et al in [3]. This method starts with a box that is big enough to contain the complete final 3D object. From the top of this box layer by layer a surplus volume is removed. They search for void regions in the upper level and corresponding bottom faces below them. The bottoms can be flat, sloppy or curvature. The subtracted volume is constructed from these tops and bottoms. Solid 3D object is constructed only by subtracting sub-volumes.

Wang and Latif developed method [24], which uses fuzzy logic to identify structural elements and to choice a suitable procedure for their reconstruction. The base principle is actually deciding whether to use the profile for extrusion or revolution. 3D solid model is then build using Boolean volumetric operations.

Table 1. Comparison of methods for reconstruction of 3D models

Year	Authors	Reference	Surface		Revolutions	Interaction	Representation		Error correction	Dimensions	Views count	Cuts	Partial views	More than 1 solution
			Planar	Cylindrical			B-Rep	CSG						
1973	Idesawa	[13]	x				x				3			
1983	Aldefeld	[1]	x	x				x			3			
	Sakurai	[22]	x	x		x	x				3			x
1992	Dutta	[7]	x	x			x				2			x
	Kondo	[15]	x				x				3			
1994	Numao	[20]	x	x			x				3			
1996	Dori	[6]	x	x			x	x		x	3			
	Itoh	[14]	x	x			x				3			
	You	[28]	x	x			x		x		3+			
1997	Masuda	[19]	x	x			x		x		3			x
1998	Shin	[23]	x	x			x				3			
	Watanabe	[25]	x	x			x		x		3			
	You	[27]	x	x				x		x	2+		x	
2001	Fujita	[9]	x	x				x			3			
	Liu	[18]	x	x			x				3			
2002	Geng	[10]	x	x		x		x			3			
2004	Cicek	[5]	x	x				x			3			
	Chovan	[4]	x				x		x		3			
2005	Lee	[16]	x	x	x			x			2-3			
	Liu	[17]	x	x	x			x			3	x	x	
2006	Benkovsky	[2]	x		x			x	x		1			
	Pavelek	[21]	x				x		x		3			
2007	Cayiroglu	[3]	x	x				x			3			
	Golovin	[11]	x	x			x	x			3			
	Wang	[24]	x	x	x			x			2-3	x		
2009	Xie	[26]	x	x				x			3	x		

The latest contribution to the solution of the problem is the method presented by Xie et al in article [26]. They focused on the recognition of hatched cuts and intersections of cylindrical hollow objects.

4. Other Methods and Analysis

Flasinski in his article [8] elaborated in detail the use of graphs to represent parts. He has described and analyzed the possible application of graph grammars to capture all the properties of objects defined by B-Rep of CSG method.

Dori and Weiss [6] address the possibility of processing the real drawings that may contain inaccurately drawn lines and dimensions. For each view a graph of restrictive criteria is created. The criteria are based on recognized dimensions. They create a network that captures all the geometric, structural and topological relationships between different entities. 3D object is reconstructed on the basis of processing of such graphs.

Golovin and Veselov [11] highlighted the fact that almost all currently available methods are still in the theoretical and test phase. Applications have been implemented to process the input projections and establish a 3D model, but often lack the link with the real CAD/CAM system. Therefore, they implemented a few modules that in addition to reconstructing the 3D model can handle DXF files, and export the resulting model to industry standard file formats IGES and STEP.

5. Patents

Fujita registered a patent [9] to convert two-dimensional CAD drawing to three-dimensional CAD drawings. Basically it is a method of utilizing CSG approach to build a solid model of volume elements generated by sweeping a profile.
Kondo et al use B-Rep to describe the solid model in patent [15]. The 2D projections are transformed to 3D space and then the cuts planes are examined in place where the profile changes. Based on this information they calculate the location of edges and surfaces, which produce a 3D model.
Numao et al have registered a patent based on the method published in article [20].

4. Conclusion

The process of automatic reconstruction of complex 3D objects is now pretty well solved by several researchers. In new solutions, it will be necessary to focus on elements that were previously excluded from processing - auxiliary and partial views, cuts and hatched areas, dimensions and geometric tolerances (as seen in Table 1).
The design of modules for CAD/CAM systems and stand-alone applications must meet the functional requirements and requirements for simple end intuitive control [12].

Acknowledgements

This article was created as a part of the project VEGA 1/0208/08 with name "3D model generating by means of reconstruction from part engineering drawing orthogonal views ".

References

1. Aldefeld, B. On automatic recognition of 3-D structures from 2-D representations. In: *Computer-Aided Design*, Volume 15, Number 2, pp. 59-64. Great Britain: Elsevier Science Ltd, 1983. ISSN 0010-4485.
2. Benkovsky, M. *Generovanie 3D modelu rotacnej suciastky*: Thesis. 64 p. KAIA MTF STU Trnava, 2006.
3. Cayiroglu, I.; Cavusoglu, A.; Celik, V. A New Method for Extracting 3D Solid Models of Objects using 2D Technical Drawings. In: *Mathematical and Computational Applications*, Volume 12, Number 1, pp. 31-40. Turecko: Association for Scientific Research, 2007. ISSN 1300-686X.
4. Chovam, A. *Generovanie 3D modelu z vektoroveho vykresu jednoduchej hranatej suciastky*: Thesis. 84 p. KAIA MTF STU Trnava, 2004.
5. Cicek, A.; Gulesin, M. Reconstruction of 3D models from 2D orthographic views using solid extrusion and revolution. In*: Journal of Materials Processing Technology*, Volume 152, Number 3, pp. 291-298. Amsterdam: Elsevier B. V., 2004. ISSN 0924-0136.
6. Dori, D.; Weiss, M. A Scheme for 3D Object Reconstruction from Dimensioned Orthographic Views. In: *Engineering Applications of Artificial Intelligence*, Volume 9, Number 1, pp. 53-64. Great Britain: Elsevier Science Ltd, 1996. ISSN 0952-1976.
7. Dutta, D.; Srinivas, Y. L. Reconstructing curved solids from two polygonal orthographic views. In: *Computer-Aided Design*, Volume 24, Number 3, pp. 149-159. Great Britain: Elsevier Science Ltd, 1992. ISSN 0010-4485.
8. Flasinski, M. Use of graph grammars for the description of mechamical parts. In: *Computer-Aided Design*, Volume 27, Number 4, pp. 403-433. Great Britain: Elsevier Science Ltd, 1995. ISSN 0010-4485.
9. Fujita, S. *Three-dimensional CAD system and method of converting two-dimensional CAD drawings to three-dimensional CAD drawing*. United States Patent US6215493B1. 2001.
10. Geng, W.; Wang, J.; Zhang, Y. Embedding visual cognition in 3D reconstruction from multi-view engineering drawings. In: *Computer-Aided Design*, Volume 34, Number 4, pp. 321-336. Oxford: Elsevier, 2002. ISSN 0010-4485.
11. Golovin, S. I.; Veselov, N. A. Automatic reconstruction of curved solids from three orthographic projections. In: *Proceedings of SYRCoSE 2007*, Volume 2, pp. 53-58. Moskva: Institute for System Programming of RAS, 2007.
12. Halenar, I.; Halenar, R. Application design and its impact on software usability. In: *Manufacturing Engineering 3/2005*, Kosice. pp. 68-69. ISSN 1335-7972.
13. Idesawa, M. A System to Generate a Solid Figure from Three View. In: *Bulletin of JSME*, Volume 16, Number 92, pp. 216-225. Tokyo: The Japan Society of Mechanical Engineers, 1973. ISSN 0021-3764.
14. Itoh, K.; Suzuki, S. Model construction from orthographic views as pseudo boolean constraint satisfaction problem. In: *Proceedings of The World Conference on Integrated Design and Process Technology*, Volume 3, pp. 76-83. Birmingham: The Society for Design and Process Science, 1996.

15. Kondo, S.; Shimizu, K.; Sato, Y. *Three-dimensional modelling apparatus and method*. European Patent EP0474458B1. 1992.
16. Lee, H.; Han, S. Reconstruction of 3D interacting solids of revolution from 2D orthographic views In: *Computer-Aided Design*, Volume 37, Number 13, pp. 1388-1398. Oxford: Elsevier, 2005. ISSN 0010-4485.
17. Liu, J.; Ye, B. New method of 3D reconstruction from mechanical engineering drawings based on engineering sematnics understanding. In: *International Conference GraphiCon' 2005*. Novosibirsk: Akademgorodok, 2005.
18. Liu, S. X. et al. Reconstruction of curved solids from engineering drawings. In: *Computer-Aided Design*, Volume 33, Number 14, pp. 1059-1072. Great Britain: Elsevier Science Ltd, 2001. ISSN 0010-4485.
19. Masuda, H.; Numao, M. A cell-based approach for generating solid objects from orthographic projections. In: *Computer-Aided Design*, Volume 29, Number 3, pp. 177-187. Oxford: Elsevier, 1997. ISSN 0010-4485.
20. Numao, M.; Masuda, H.; Shimizu, S. *Solid model construction method and apparatus*. European patent EP0592126B1. 1994.
21. Pavelek, I. *Generovanie stenoveho 3D modelu jednoducheho hranateho telesa z vektoroveho zaznamu priemetov*: Thesis. 84 p. KAIA MTF STU Trnava, 2006.
22. Sakurai, H.; Gossard, D.C. Solid model input through orthographic views. In: *ACM SIGGRAPH Computer Graphics*, Volume 17, Number 3, pp. 243-252. Detroit: ACM, 1983. ISSN 0097-8930.
23. Shin, B. S.; Shin, Y. G. Fast 3D solid model reconstruction from orthographic views. In: *Computer-Aided Design*, Volume 30, Number 1, pp. 63-76. Great Britain: Elsevier Science Ltd, 1998. ISSN 0010-4485.
24. Wang, Z.; Latif, M. Reconstruction of 3D Solid Models Using Fuzzy Logic Recognition. In: *Proceedings of the World Congres on Engineering 2007*, Volume 1, pp. 37-42. London: Newswood Limited, 2007. ISBN 978-988-98671-5-7.
25. Watanabe, T. Revision of Inconsistent Orthographic Views. In: *Journal for Geometry and Graphics*, Volume 2, Number 1, pp. 45-53. Lemgo: Heldermann Verlag, 1998. ISSN 1433-8157.
26. Xie, B.; Chen, L.; Chen, L. A new reconstruction method based on the volume-based approach. In: *Proceedings of the 2009 International Symposium on Information Processing*. pp. 132-135. Huangshan, China: Academy publisher, 2009. ISBN 978-952-5726-02-2.
27. You, C. F.; Yang, S. S. Automatic Feature Recognition from Engineering Drawings. In: *The International Journal of Advanced Manufacturing Technology*, Volume 14, Number 7, pp. 495-507. London: Springer-Verlag, 1998. ISSN 0268-3768.
28. You, C. F.; Yang, S. S. Reconstruction of curvilinear manifold objects from orthographic views. In: *Computers & Graphics*, Volume 20, Number 2, pp. 275-293. Great Britain: Elsevier Science Ltd, 1996. ISSN 0097-8493.

.

Using a Function of Image's Informativeness for Analyzing its Spatial Structure

Y. Vichevskaya
Izhevsk State Technical University
Izhevsk, Russia
e-mail: ulka_him@mail.ru

A. Murynov
Izhevsk State Technical University
Izhevsk, Russia
e-mail: po@istu.ru

Abstract[1]

Powerful visualization, spatial representation of objects and events, a wide range of special analyzing and modeling programs create for a user a comfortable environment for solving the intellectual problems.
The process of transfer of graphic information in electronic form consists of two phases - scanning and recognition. Modern methods of manual and automatic recognition of images do not allow to process large volumes of data.
To achieve a high quality of recognition we need a new method for analyzing the structure of graphic documents. Based on the function of informativeness, this method will allow to analyze the structure of images and ensure high performance in creating the electronic documents.
The idea of a function of informativeness gives an opportunity to describe a formal ways (algorithms) of finding geometric features on images. These algorithms will be used for writing programs that automatically generate a language for describing images.

1. Introduction

It is well known that over 90% of the information people receive through the vision. That's why the technical progress affects primarily the means of collecting and processing visual information. The development of means and methods for processing, analyzing and image recognition in recent decades has shown that the capability of extraction useful information from images is determined by the properties and characteristics of their spatial structure.

2. The Need for a New Method of Image Recognition

One of the basic methods of image processing is the recognition of graphic documents [2]. Today there are a lot of methods of image recognition. But applying these methods does not always ensure qualitative result. Therefore, actual direction is development of a new method for recognizing graphic documents, based on the idea of the function of image's informativeness, which simulates the reactions of the human visual analyzer. This idea was developed by I.B. Muchnik and N.V. Zavalishin [1].
When a human looks at any image, by combining the various transformations, he almost always can find some features on it. Location of the gaze fixation's points is determined by geometric structure of analyzed image, and fixation points are concentrated on areas of the image, which might be called local geometric features of the image (a break, the end of a line, cross, etc., Fig. 1).

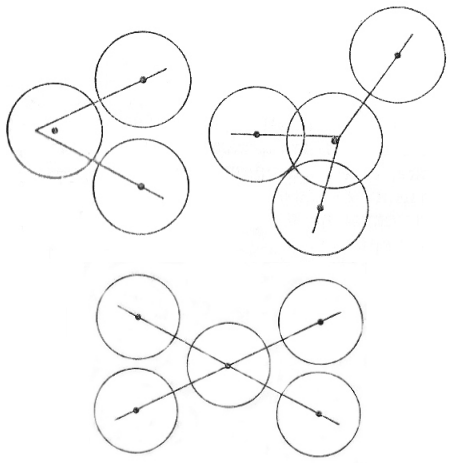


Figure 1. Location of the gaze fixation's points on the geometric features of the image (a break, the end of a line, cross)

The important task is development a formal method of finding geometric features, which has not relied on pre-defined list of types of fragments' forms. A natural way of developing this method consists of dividing the process into two phases: finding the geometric features, without analyzing or indicating the type of fragments'

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

form, and recognition or classification of found fragments' form.

3. A Method of Constructing a Function of Informativeness

To select the local fragments of the image we need to use a "window", which size is smaller then the size of the image. Assuming that the "window" is placed on the image, it will select some fragment. It's necessary to assess the degree of image's "regularity" within the window, and receive a number as a result of this assessment, which will be ascribed to the point of the image's plane, that will coincide with the center of the window. By placing the center of the window at each point of the image plane and evaluating the degree of variation of the image's "regularity" within the window, we get a function. Its extrema are located just on those areas which are the geometric features. After finding the extrema, there is recognition of a specific form of each of the found geometric features. The fragments of the image, which have the extrema, are called informative, and a function is a function of informativeness.

In order to describe the degree of the image's "regularity" within each fragment the standard image is introduced into consideration. It's a spot, whose size coincides with the size of the window, and blackout falls from center to periphery. The standard image, in essence, plays the role of the standard fragment with the greatest possible irregularity - a black spot on a white background (Fig. 2).

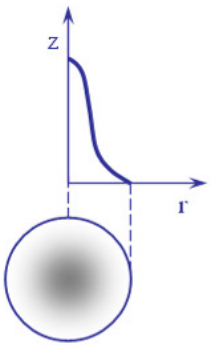


Figure 2. The standard image

As the number characterizing the degree of "regularity" or "complexity" of the image within the current fragment a measure of differences between the fragment and the standard image is considered, understood as the distance between these images. To determine the distance it's necessary to set space. This space is the space of blackout's values in the cells of the window. It is called the space of blackout's vectors.

The calculation of a function of informativeness is determined by the formula:

$$I(t,s) = I_B(t,s) - 2I_A(t,s) \qquad (1)$$

I - value of a function of informativeness,
t, s - offset relative to the center.

$$I_A(t,s) = \iint_{A(t,s)} F(x-t, y-s)dxdy \qquad (2)$$

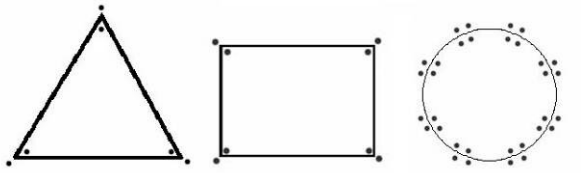
$$I_B(t,s) = \iint_{B(t,s)} F(x-t, y-s)dxdy \qquad (3)$$

F - original image,
A - area of a standard image,
B - area of the window,
x, y - coordinates of the window's center.

The function of informativeness takes the smallest value on the fragments, which blackout is distributed the same way as a blackout of the standard image. In other words, the extrema of function of informativeness should be placed in locations of significant blackout's variation of the image. At large monochrome areas a function of informativeness doesn't have the extrema, because when the window is moving within these areas it cuts a fragment, that does not change.

4. Geometrical Characteristics of the Local Extrema of a Function of Informativeness

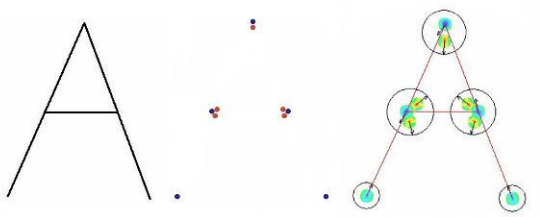
On the contour images the locations of the greatest blackout's variation are the geometric features such as line ends, corners, branches, etc. The Fig. 3 shows the result of finding the extrema of a function of informativeness of simple geometric objects.



a **b** **c**

Figure 3. Searching the extrema of a function of informativeness of simple geometric objects: a) triangle, b) rectangle, c) circle

The function of informativeness makes it possible to select as characteristic almost all types of fragments corresponding to the geometric features of complex images. In addition, the characteristic fragments are selected uniformly and centered, regardless of "context" of images. On the Fig 4 there is the result of finding the extrema of a function of informativeness of the letter "A". In the future the structure of the letter can be determined based on the configurations of the extrema: the minimums of the function of informativeness are at the ends of lines and external corners, and the maximums are for the inside corners (Fig. 4 c).



a **b** **c**

Figure 4. Searching the extrema of a function of informativeness of the letter "A": a) original image, b) the extrema of a function of informativeness, c) determining the structure of the letter with the found extrema

5. Conclusion

Calculating the function of informativeness can create a structural description of the image that will be the basis for recognition of images.

Thus, analysis of the configuration of the local extrema of a function of image's informativeness can be the basis of a new method for analyzing the graphic documents. It's necessary to select the analysis of texts, images and complex graphic documents (maps, drawings, charts) among the directions of realization of the method.

References

1. Zavalishin N.V, Muchnik I.B. "Visual Perception Models and Image Analysis Algorithms". Moscow: Nauka, 1974.
2. Pratt W. "Digital image processing". Moscow: Mir, 1982.

Water resources management based on service-oriented approach in geographic information systems

S.V. Pavlov
Geographic Information Systems Department
Ufa State Aviation Technical University
Ufa, Russia
e-mail: psvgis@mail.ru

R.A. Shkundina
Geographic Information Systems Department
Ufa State Aviation Technical University
Ufa, Russia
e-mail: r.shkundina@mail.ru

Abstract[1]

In the paper we analyze different types of physical architectures of the geographic information systems on the sample of the development of corporative geographic information system Rosvodresursy. The transition from desktop to internet application based on the service-oriented platform of the ArcGIS is described.

1. Introduction

Water as a variety of natural resources play an important role in the development of Russian economy, so the rational management of water resources is one of the primary tasks of life-sustaining state. Water resources management in Russia is controlled by the Ministry of Nature and Environment and the Federal Agency of Water Resources (Rosvodresursy) [5]. Because of the spatial extent and distribution of water facilities throughout the Russian Federation geographic information system technology is proposed to be used. Therefore, for the automation of many tasks in water resources management we develop corporate geographic information system for different levels of management, which main goal is to provide a GIS user with the consistent, reliable, relevant and accurate spatial, reference and analytical information for water resources assessment and decision-support of operational and strategic decisions on water management in Russian Federation.

2. Thick-client architecture of GIS

Now GIS Rosvodresursy covers all levels of management (federal, basin and territorial) and based on "thin"-client architecture [4]. It contains several subsystems: subsystem of modelling pollutants, spills and distribution of oil and oil products in water bodies and the threat of their falling into the water bodies; subsystem of complex analysis of water resources in Russia; subsystem of zoning Russian Federation territory by the range of indicators that characterize the impact and the state of water objects; subsystem of flood zones' modeling for the determination of places most vulnerable to flooding during the flood period, an assessment of the risk associated with the floods; subsystem of reporting the results of modelling; subsystem of reference and search tasks; subsystem of federal reservoirs; subsystem of integration with information system of analyzing critical situations , etc.
In the beginning software architecture of the GIS Rosvodresursy was a client-server with "thick" client (fig.2, designations are on the fig.1) based on the ESRI ArcGIS Desktop software [1]. A user gains an access to the geospatial and attributive information located in the databases through ArcSDE (SDE – Spatial Database Engine), which represents server software for the storage and management of spatial data in relational database systems. In the systems within the "thick" client architecture all actions with data: processing, generation of graphical information, modeling, etc. is executed on the client-side. Therefore such architecture requires high computing power of client workstations. So the following components are used in the first version of the GIS Rosvodresursy:

- standard functions of the ArcMap application, e.g. visualization of spatial and attributive data in the GIS Rosvodresursy;
- custom geoprocessing models of the ArcToolbox, e.g. river flood modeling system;
- special .NET program modules for the ArcMap application, e.g. water resources spatial classification subsystem.

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

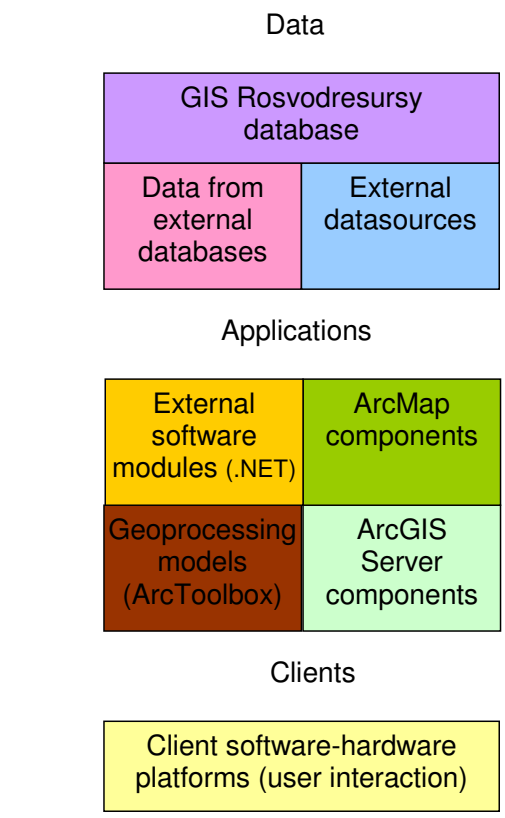


Figure 1. Types designations of the GIS Rosvodresursy software components

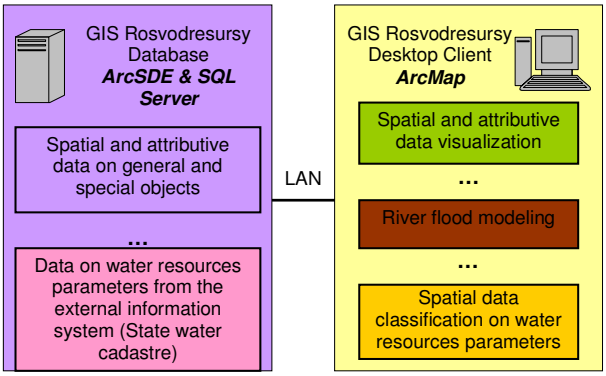


Figure 2. Client-server architecture of the GIS Rosvodresursy with "thick" client

The restrictions of the thick-client architectures of geographic information systems are:

° limited cross-system mobility of the GIS subsystems;
° high requirements to user workstations software to implement data processing;
° difficulties of installation and support of the "thick" clients for the geographically distributed users;
° high requirements to user workstations hardware due to extensive data processing on the client side;
° tricky implementation of the interaction with other enterprise information systems due to lack of common interfaces, every system requires special interface module to obtain subject data, for example on the hydraulic structures or water monitoring stations;
° additionally, subject data collection and actualization requires using of mobile devices with GPS or GLONASS satellite navigation feature.

So, the thick-client architecture of GIS Rosvodresursy implies several shortcomings which are inherent to "thick" client software approach and must be solved.

3. Service-oriented architecture of GIS

The most promising directions for the development of corporate GIS to overcome client-server architecture shortcomings is implementation of service-oriented technologies, mobile GIS components and using OLAP to enhance integration of GIS with other information systems of the organization [2,3].

In the framework of the service-oriented architecture common data exchange infrastructure for the Web-service software modules is created. So the cross-system integration is simplified, software and hardware requirements are eased, supportability of the system and reusability of the components are increased. Software module being implemented as a web-service can be used by different types of clients through standard data exchange interface. All this abilities are provided by ESRI ArcGIS Server software [1]. New technology implementation requires resolving of the task of existing software reusing and system components integration, including GIS components. Software modules physical partition and interaction in the service-oriented architecture is a result of steady development of the previous architectures (component, client-server).

Aforementioned desktop GIS Rosvodresursy was gradually transformed into the Internet application (Web-site) functionally comparable with desktop application [2]. One of the main tasks of this transformation was relocation of the software modules (.NET and geoprocessing) from the client to the server side, keeping intact main algorithm and redesigning interface. The client side of the application implemented on the Flex technology, server side implemented as ArcGIS Server GIS-services and custom .NET modules (fig.4). So the most part of the tasks from the desktop client implemented now on the server side as a set of Web-services. Intermediate step of the Web-application development was based on the ArcIMS or ASP.NET platform. In such implementation most tasks were already relocated to the server side as Web-application (fig.3) resolving three problems of the desktop client-based GIS. But problems of cross-

system mobility and interaction still exist due to inherited module partition and interaction scheme.

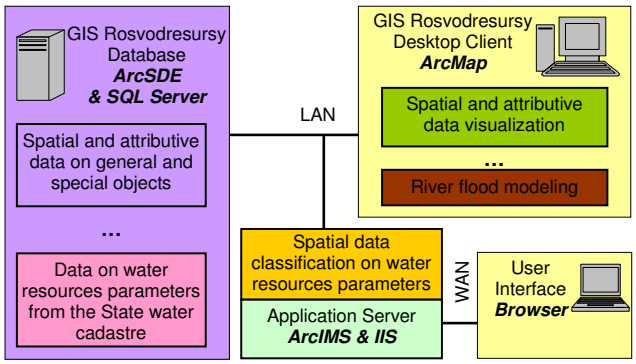


Figure 3. Client-server architecture of the GIS Rosvodresursy with "thick" and "thin" clients

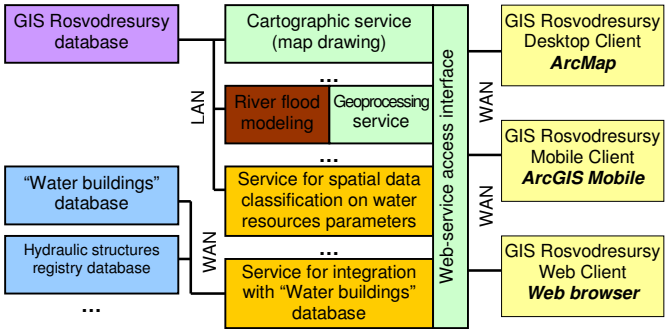


Figure 4. Service-oriented architecture of the GIS Rosvodresursy

ArcGIS platform provides to the developers number of capabilities in implementing service-oriented technology in GIS (utilization of some of them is shown further on the GIS Rosvodresursy examples):
1.GIS-services (i.e. Web-services implementing GIS tasks) of the ArcGIS Server – provide maps, spatial data, spatial models and tools to the different types of clients.
2.Software extension of capabilities of the ArcGIS GIS-services by custom .NET modules.
3.ArcObjects components for custom Web-services creation.
Use of these ArcGIS Server technologies gives the ability to implement in the Web-application some features, previously available only for ArcGIS Desktop application users, and provides access to the same service-implemented tasks for different types of client software. For example GIS Rosvodresursy includes: spatial data editing subsystem, resource-demanding tasks of modeling of floods and contamination propagation, water basins' building, etc.

1. ontclair State University, USA, 2001.

Special objects spatial and attributive data editing subsystem available in the GIS Rosvodresursyy as Web-application, as well as mobile application. Data input and editing Web-application of the GIS Rosvodresursy gives the ability to input and modify subject objects location without special software for the users from central office of Federal Agency of water Resources and its local branches which have more actual information. Using GIS Rosvodresursy mobile client software users can add, delete and move objects on the electronic map, for example to specify location of the water consumers, monitoring stations, hydraulic constructions.
Flood modeling subsystem based on the water level measurement data is an example of the tasks previously available only for desktop GIS (ArcGIS Desktop) users. Modeling subsystem was redesigned into ArcToolbox geoprocessing model and published as a geoprocessing service giving access to this task for the GIS Rosvodresursyy Web-application users by means of the ArcGIS Flex API controls.

4. Conclusion

So, the implementation GIS Rosvodresursy with service-oriented technology resolves some problems specific for desktop systems: installation and support difficulties, high demand to hardware and software, complicated integration with another information systems and increases efficiency of the water management tasks: data exchange between different branches and the central office, access to the dynamic data about water objects state, hydraulic structures, water users, integration with other information systems of the Rosvodresursy.

References

1. ESRI Corporation website: http://www.esri.com
2. Geospatial Service-Oriented Architecture (SOA). An ESRI White Paper. June 2007.
3. Ivanov I.G., Pavlov S.V., Nikitin A.B. "Spatial data exchange in the distributed GIS of the Rosvodresursy based on the ArcGIS Server". ArcReview, №4 (47), Data+, Moscow, 2008, pp.2-4.
4. Ivanov I.G .Pavlov, S.V., etc. Organization of spatial data sharing in distributed GIS Rosvodresursy based on ArcGIS Server. ArcReview № 4 (47) 2008.
5. Nikitin A.B., Pavlov S.V., etc. "Geoinformation system of Federal agency of water resources". ArcReview 2006; 1:6-7

Virtual Firefighting Simulator

Nils Schmeißer
Information Technology
FZ Dresden-Rossendorf
Dresden, Germany
e-mail: n.schmeisser@fzd.de

Abstract[1]

A simulator will be presented which employs VR techniques in order to provide a training environment for fire-fighters. While existing systems for fire-brigade training do focus on tactics and command training, this simulator targets the training of the fire attack group. It provides scenarios for search & rescue operations, direct firefighting using water and foam as well as radiation source search operations.
Visual, audio and haptic feed-back is provided. The system supports immersive displays (HMD) as well as standard techniques (monoscopic big-wall). It implements a walking paradigm, which enables the trainee to move around freely.
The simulator uses extended reality features in order to improve the handling (nozzles, dose meter, ...). The trainee can use the system fully equipped, even using the respiratory protection, in order to provide a realistic feeling.
Several physical models are used to provide effects closed to reality, namely:

- heat release rate is used to model the energetic behavior of a fire, flame size etc.
- a simple energetic model for fire-extinction by cooling
- water and foam ejection (depending on pressure, volume, form of jet, agent)
- a smoke system (visibility equivalent to height of eye)
- superposition of radiation sources (no shielding) and detection by Teleprobe FH40G

The simulation is supervised by an instructor, each action, e.g. radio traffic, movements is logged for analysis.

1. Introduction

Firefighter and other rescue personal needs to be trained regularly in order to react properly in complex situations. The very important aspect of collaborative operation between different personal can almost be trained in virtual environments [1, 2, 3, 4].
For the individual member of a rescue team, e.g. the attack crew a fully immersive training environment is needed to train special actions like fire extinguishing, movements and tactics [10]. These environments can be set up as scaled real world environments [5]. They are expensive and, although they are quite save, do adhere threads to the health of the user.
These environments could be transformed into virtual reality environments if the appropriate system follows the main "principle of space and cyberspace" as described in [6]. These principles map real world experience to VR concepts.
Nether the less these real world training facilities can not be fully replaced by pure VR systems as they do provide experiences which can not be simulated right now (heat, free physical movement, mechanical limitations while handling hoses and nozzles, ...).

2. Methods

As proposed by Zeltzer [11] the concept of *presence*, *interaction* and *autonomy* lead to immersive VR environments.

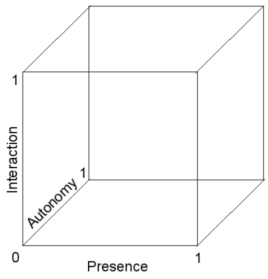


Figure 1. Zeltzer Cube

2.1. Presence

The experience of the virtual world should be closed to reality. Therefore a system should produce data for vision, hearing, touch, balance & kinesthesia, taste, smell.

[1]**Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

2.2. Interaction

Firefighter interact using several devices like nozzles and radio. These devices should provoke appropriate changes in the simulation.

2.3. Autonomy

The state of the system should strictly follow physical rules independently. In most real world fire training setups an operator decides if a fire is extinguished manually!
Recent advances in the correct physical simulation of fire, fire spreading and evolution of fire provide the basis for the computer simulation of hazardous scenario [8].
The application of these concepts for a firefighting simulator is discussed in the following chapters.

3. The Firefighting Simulator

3.1. System overview

The system presented here provides a specialized user interface and standard output technologies. The simulation is driven by a simulation-engine which can provide hazardous scenario (fire, explosion, injured person, radioactive sources).

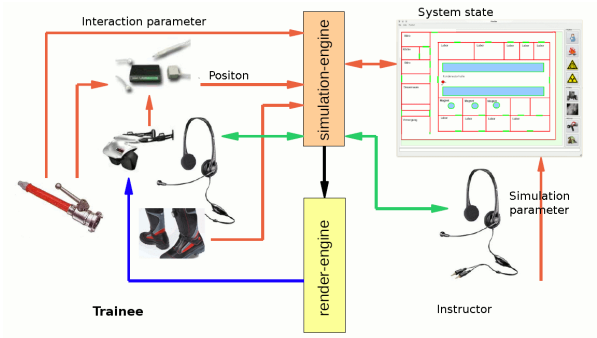


Figure 2. System overview

The trainee is equipped with a tracking system to estimated her/his position and orientation in the virtual world. A head mounted display is used to provide visual perception. The display is driven by a render engine which loads static geometries via VRML formatted files. The render-engine provides functions for visualizing fire, smoke, explosions and persons.

3.2. Render-engine and display system

The render engine is simple OpenGL application which runs standalone and receives visualization parameters via the network. It can produce monoscopic and stereoscopic (anaglyph, dual-head stereo, frame sequential stereo) output.

Fire (flames) are modeled as multi-particle systems. Smoke is displayed employing the glFog capability combined with a horizontal, transparent top-plane.
Currently either a head mounted display (eMagin Z800) or a big screen rear-projection with active stereo (Xpand IR shutter glasses) is used. While the big screen projection could be used by a crew, the HMD provides full immersion into the training environment.
For fire engagement scenario a stereoscopic output is usually not necessary as in most cases the crew will experience dense smoke with rare visibility conditions.
Combined with a set of speakers the trainee will perceive audio-visual output (explosions, crying people, ...). All other senses are not touched, as the appropriate output hardware is not available or too expensive. Taste and smell can be disregarded as crew is usually equipped with respiratory protection system.

3.3. User input

In order to compute correct visualization parameters the simulation must know about the position of the trainee. Therefore the trainee is attached to a head tracking system (polhemus Patriot). Although the tracker allows to find the relative position this is not sufficient to enable the trainee to perform a walk-through in a natural manner. As virtual floors are still subject to research (Cybercarpet) or too expensive (Virtusphere) this technology is not available here. It can be approximated by a “walk-in-place” paradigm. The trainee is attached to a step-sensor which allows movement in the viewing direction.
As different scenario require different devices a set of several virtual devices is available (e.g. a C-size nozzle and a Teleprobe FH40G dummy).

3.4. Simulation-engine and scenario

Prior to start, all parameters of an engagement are defined by an instructor. A scenario can be a combination of different basic scenarios. The instructor can change the scenario any time during the simulation. The operation of the simulation-engine will be discussed by example.

3.4.1. Search-and-Rescue operation

Objective: one or more missing person need to be located (and rescued).
First responders usually assigned the task of exploring the scene and perform basic rescue operation, mainly rescuing (disabled/insured) people from the hazardous environment. They therefore need to train communication skills and procedures of exploration.
The system described here emulates radio-communication via head-sets attached to a multi-channel sound-card. All radio-traffic is logged and can be analyzed later.
Tactics for search and rescue procedures vary on different environmental parameters like (building)

geometry, including dynamic geometries, visibility (smoke), heat and others.

Free movement and proper visual presentation is required to train proper response. The system uses a polhemus Patriot tracking system [12], which provides stable location and orientation data. These data are transformed into a position in the virtual environment. As the detection range of the tracking system is limited and the user is bound to a certain area because of cabling, free walk is possible only in a very limited space. To break this limitation a virtual walking paradigm is implemented, using a stepping-sensor (simple switch) which is attached to the foot-wear of the trainee. Each "step in place" is interpreted as a step into the direction of the current view. While the height of eye and orientation of the trainee is provided by the tracker, "free" movement within the virtual world is possible.

Figure 3. Tracking sensors

In order to provide stress to the trainee (as she or he will experience in real environments) one or more casualties are placed in the scenery.

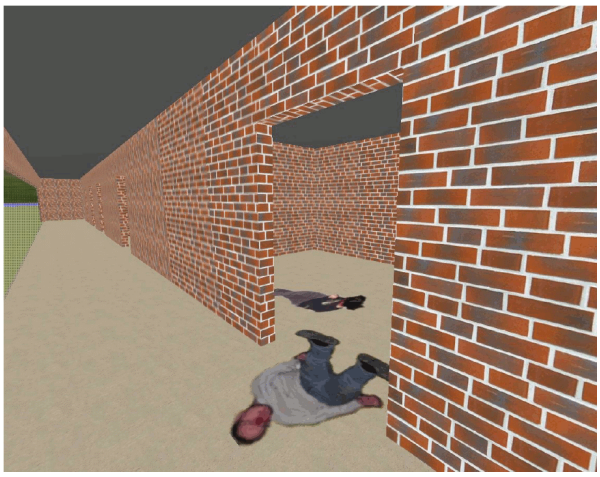


Figure 4. Search & Rescue Mission view

Depending on the state (in danger, escaping, immobile, rescued) of the casualty, its grade of injury (not insured, light, medium, heavy) and the vital state (NAD, unconscious, apnea, cardiac arrest) different presentations for visual and audio perception (cries, moaning) are generated. The perceived sound level represents a superposition of all currently placed casualties.

3.4.2. Radiation protection

Objective: estimate activity, locate one or more radiation sources.

Radioactive materials are widely used in medicine, industry and research. Therefore rescue workers have to prepare for scenarios involving radioactive materials. The main tasks consist of:

1. determining contaminated area
2. rescue people from "dirty area"
3. determine grade of contamination of persons
4. prevent other areas from contamination
5. locating radiation sources

To perform these operations firefighters are equipped with activity sensors (e.g. Teleprobe FH40G). Training the use of these devices usually requires real (but small) test sources. These source have to be supervised by specialized personal. Which makes training expensive. As physical laws of measurement are well known (and quite simple) they can be implemented very easy in a virtual environment. This system handles a set of radiation sources which are parameterized by the isotope and activity. In Germany rules of engagement are given by [13]. Three types of detection are required: dose rate for localizing "dirty" areas and estimation of threads, doses received and proof of contamination (proof of activity). While activity is directly correlated to the activity of the source, dose rate and dose are measured indirectly. To simulate the effect of dose rate, the so called "dose conversion factor" *DCF* can be used to compute the dose rate *H* from the activity *A*

$$H = DCF \cdot A \qquad (1)$$

Doing a superposition on all sources in the virtual environment the dose rate at a certain point x_p can be estimated by

$$H_p = \sum DCF_i \cdot A_i \cdot \frac{1}{\left\| \underline{x}_i - \underline{x}_p \right\|^2} \qquad (2)$$

where x_i is the position of source. As the geometry and materials, the virtual world consists of, is known, damping could be taken into account (not implemented currently). Because the Teleprobe FH40G does not have directional characteristics, the estimated value represents the one which would have been measured in a real environment.

The computed value (H_P) is mapped to the virtual display of the Teleprobe FH40G. The user controls the location and orientation of the Teleprobe in the virtual world using a dummy with a second scan-head attached.

3.4.3. Firefighting

Objective: locate fire, fight fire

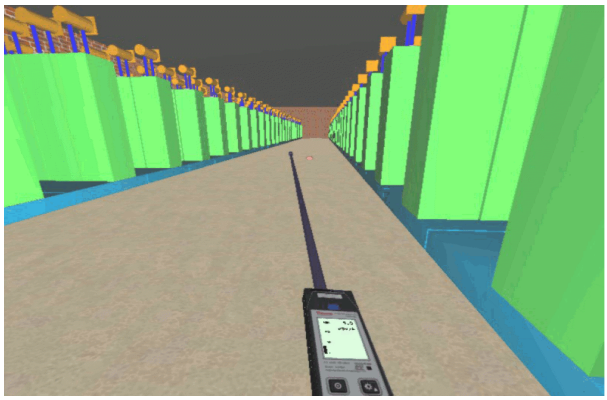


Figure 5. Doing exploration in a "dirty" environment

Fighting a fire itself is considered one of the main tasks of fire brigades. To perform this task well a lot of experience is required. These skills have to be trained regularly. Nowadays real world training is performed using wood or gas heated fire houses (Fig. 6). These installations allow training under realistic but controllable conditions.

Figure 6. Fire-house training (FF Rudersdorf, Austria)

As knowledge about the physical and chemical processes of fire extends it is almost possible to simulate standard situations, including flash-over and back-draft.
Modeling the behavior of fire is a complex task. Currently most models employ continuous fluid dynamics methods on compressible media to calculate the combustion process [8, 14, 15]. These methods are very advanced but require a lot of computational power. For a training environment it might be sufficient to reduce the complexity of the model and do a rough estimation of the combustion process employing the concept of heat release rate [9,16]. For simple materials the heat release rate *dq/dt* (HRR) can be deduced from

$$\dot{q} = \dot{m} \cdot \Delta H_c \qquad (3)$$

Where *dm/dt* is the mass loss rate and ΔH_c is the known lower heat of the material. HRR itself gives an estimate on the heat per area released at a certain time *t*. If the material is a composite the HRR could be measured by a calorimeter [17].

Figure 7. Combustion of Bunk Bed

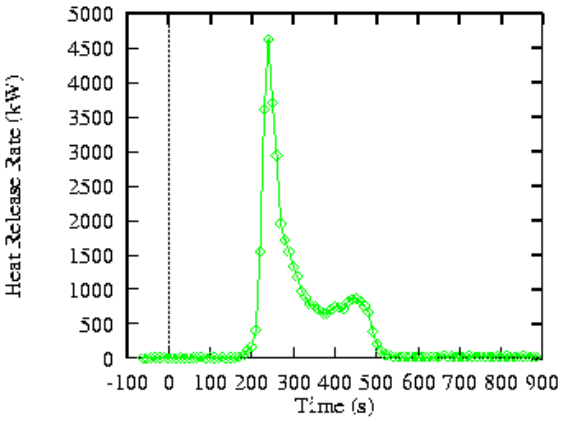


Figure 8. HRR measured for burning of bunked bed

Given a certain material, its mass, the HRR and the size of the burning surface, the heat emitted by the fire can be estimated by

$$Q_F(t) = \int \dot{q} \cdot A_F \, dt \qquad (4)$$

[18] suggests a very simple correlation between average flame hight and mass loss rate.

$$\log(h_F) = -0.38 + 0.5 \log(\dot{m}) \qquad (5)$$

As the mass loss rate can be computed from (3), all necessary parameters for creating an appropriate visual representation of a fire are available now. Flames are commonly displayed using multi-particle systems with textures. With respect of the chaotic nature of flames, the following generator is used

$$\underline{x}_i t = \begin{pmatrix} r_F \cdot md \cdot \cos(2\pi \cdot rnd) \\ r_F \cdot md \cdot \sin(2\pi \cdot rnd) \\ \vartheta \cdot md \end{pmatrix} \cdot t \qquad (6)$$

with breaking condition $x_z < h_F * rnd$. An example snapshot of the resulting fire representation is shown in figure 9.

With that model even flame drift can be visualized to improve perception.

Concept of fire extinction

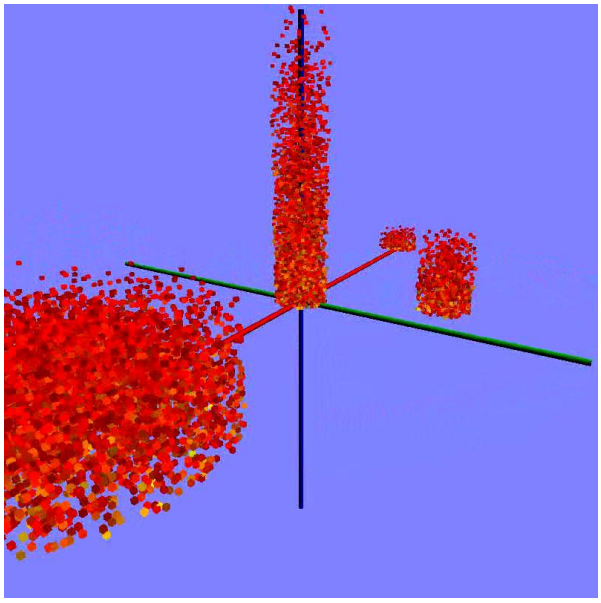


Figure 9. Flame generator example

From the combustion triangle the main concepts of fire extinction can easily be deduced:

- extinction by cooling
- extinction by smothering or replacing oxidant
- extinction by removing combustible

The process of extinction by cooling can be modeled using the following relation:

$$Q(t) = Q_F(t) - \int \dot{q}_L dt - Q_{CTR} \qquad (7)$$

Assuming the fire to be defined by its area A_F and its HRR and assuming a "flat" behavior of the extinguishing agent. q_l/dt is approximated by

$$\dot{q}_L = \frac{A_{LF}(t)}{A_L(t)} \cdot I_L(t) \cdot C_L \cdot S_L(t) \qquad (8)$$

Where A_{LF}/A_L defines the ratio of the agent applied to the fire (A_L is considered to be the area on the ground covered when the agent's jet hits the ground, A_{LF} is that part of A_L which overlaps with A_F). $I_L(t)$ gives the volume per time, C_L the heat capacity and $S_L(t)$ a measure of effectiveness of the agent (e.g. depending on the drop size). Usually $I_L(t)$ and $S_L(t)$ will have discrete values given by the extinguishing device used.

The fire will extinguish if the combustible is gone (mass loss rate) or if the temperature at the surface of the combustible falls below the ignition point of the combustible. At this point conduction, transmission and radiation of heat should be taken into account, which again leads to very complex models. Here these effects are modeled by a constant correction-term Q_{CTR}.

The temperature in question could be determined by

$$dQ = C \cdot m \cdot dT \qquad (9)$$

Where C and m are specific constants given by the combustible-air interface.

In order to find A_L and A_{LF}, the trajectory of the extinguishing agent needs to be modeled. For "heavy" agents like water, low- and medium expansion foam this can be done by modeling the jet consisting of a set of particles. The trajectory can be computed as

$$x_n(t) = \frac{m_p}{\beta} \cdot \vartheta_0 \cdot \cos\alpha \cdot \left(1 - e^{-\frac{\beta}{m_p} \cdot t}\right) \qquad (10)$$

and

$$y_n(t) = \frac{m_p}{\beta} \cdot \vartheta_0 \cdot \sin\alpha \cdot \left(1 - e^{-\frac{\beta}{m_p} \cdot t}\right) + \left(\frac{m_p}{\beta}\right)^2 \cdot g \cdot \left(1 - e^{e^{-\frac{\beta}{m_p} t}}\right) - \frac{m_p}{\beta} \cdot g \cdot t + h_0 \qquad (11)$$

Fig. 10 shows simulated trajectories with respect to the attack angle α.

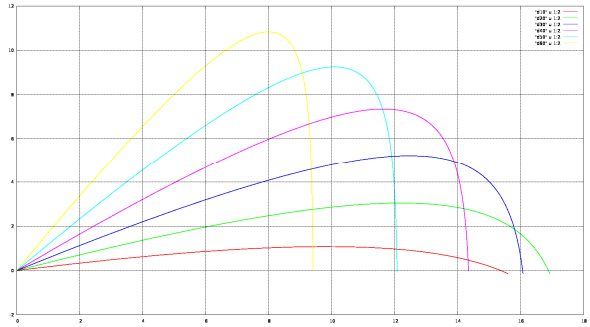


Figure 10. Trajectory examples

The air resistance β depends on the form (full beam or spray) and the expansion rate (if foam is simulated). The mass m_p is determined by the delivery pressure, the density of the agent and the cross-section of the hose

$$m_p = A_0 \sqrt{2 \cdot p \cdot \rho} \cdot dt \qquad (12)$$

Assuming no pressure loss in the fire hose, the discharge velocity v_0 can computed by

$$v_0 = \sqrt{\frac{2 \cdot p}{\rho}} \cdot \frac{A_0}{A_1} \qquad (13)$$

where A_1 is the cross-section at the nozzle.

Finally, the trajectory is mapped to a plane orthogonal to the ground, where the direction of this plane is defined by the azimuth angle $\gamma$of the nozzle.

The parameters are controlled using a modified C-size nozzle with a second polhemus scan-head attached at the end of the nozzle. The trajectory plane is translated to the absolute position of the nozzle defined by

$$\tilde{\underline{x}} = \underline{x}_{fighter} + x_n(t) \cdot \begin{pmatrix} \cos\gamma \\ \sin\gamma \\ 0 \end{pmatrix} + y_n(t) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad (14)$$

The simulation engine permanently monitors the extinction parameters and computes the process parameters in real-time. The visualization is updated in real-time too.

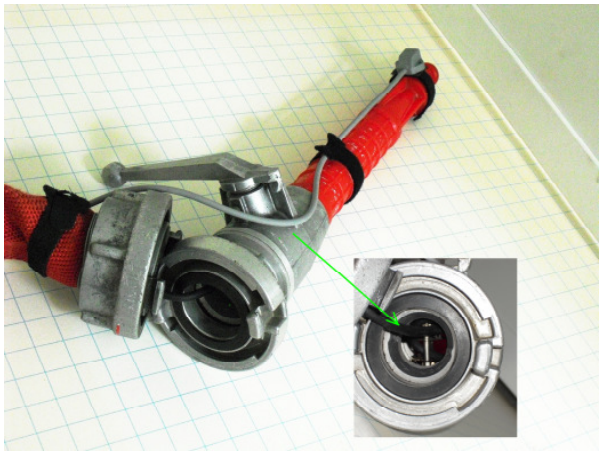


Figure 11. Modified C-size nozzle

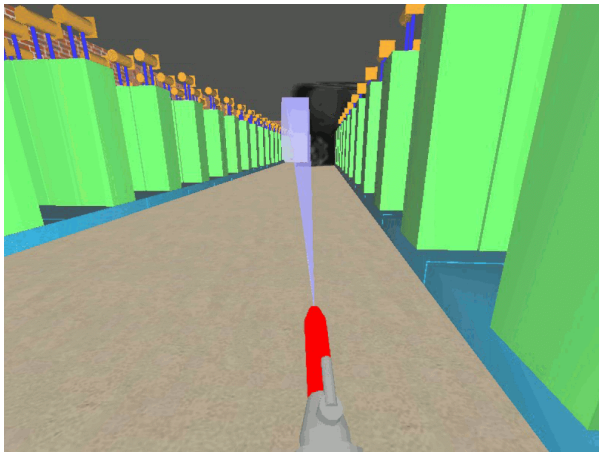


Figure 12. Firefighting using C-size standard nozzle

4. Conclusion

The system is operational and was presented on several occasions. Some tests showed, that free movement is limited if using the big wall setup. Therefore the fully immersive setup using a head mounted display should be used for "free" training, while the big wall setup could be used to train actions which do not require a walk-trough.
The implementation of the basic scenarios have been confirmed to be appropriate in presentation and behavior. The correct (physical) behavior of the processes simulated still needs to be veryfied. A first experiment showed reasonable results [19].
Future extensions will include autonomous persons and better fire simulation using the FDS.

Using the system in addition to real world training facilities – either to prepare a real world training or to train particular situations – could help to improve the abilities of firefighters.

5. Acknowledgements

The project is supported by IT (FZD) and the campus fire brigade. As part of the education program at FZD, trainees of FZD participated in mechanical and electronical interface design. Several colleagues of FZD and VKTA contributed knowledge and experience.

Figure 13. Principle usage of the simulator in a big wall setup

References

1. Lohman, S., "Simulationstechnologie für Ausbildung und Training von BOS", 5. Virtual Fires Kongress, 2010
2. e-semble, "XVR", http://www.e-semble.com/xvr/
3. VSTEP, "RescueSim", http://www.rescuesim.com
4. Astragon, "Feurwehrsimulator 2010", http://ww.astragon.de/
5. The Falk Rotterdam International Safety Center (RISC), http://www.falk.nl
6. Benedikt, M., "Cyberspace, first steps", The MIT Press, 1991
7. Kalawsky, R. S. "The Science of Virtual Reality and Virtual Environments", Addison-Wesley, 1993
8. National Institute of Standards and Technolgy, "FireDynamicsSimulator" (http://fire.nist.gov/fds)
9. Bryant, R. A. et.al., "The NIST 3 MW Quantitative Heat Release Rate Facility – Description and Procedures", NISTIR 7052, 2004
10. Brauner, Ch., "Möglichkeiten und Grenzen der Simulation in der Feuerwehrausbildung", 3. Virtual Fires Kongress, 2008
11. Zeltzer, D., "Autonomy, Interaction and Presence", Presence 1(1), 1991
12. Polhemus, http://www.polhemus.com
13. FwDV 500, Feuerwehr Dienstvorschrift 500 "Einheiten im ABC-Einsatz", 2003
14. Melek, Z., Keyser, J., "Interactive Simulation of Fire", Technical Report 2002-7-1, CS Texas A&M University, 2002
15. Ishikawa, T. et.al., "Visual Simulation of Spreading Fire", University of Tokio and Hokkaido University
16. Bryant, R.A., Mulholand, G.W., "A guide to characterizing heat release rate measurement

uncertainty for full-scale fire tests", Fire and Materials Fire Mater. 2008/32, 2008
17. NIST Building and Fire Research Laboratory, http://www.fire.nist.gov/fire/fires/fires.html
18. Weise, D.R. (USDA Forest Service), Fletcher T., Smith S. (Brigham Young University), Mahalingam, Zhou X., Sun L. (University of California), "Correlation of mass loss rate and flame height for live fuels"
19. Campus Fire Brigade, Rossendorf, https://www.fzd.de/wfw/2009-LV/

Fig. 6 Freiwillige Feuerwehr Rudersdorf (Austria)

Fig. 7 & 8: NIST Building and Fire Research Laboratory, http://www.fire.nist.gov

Interactive 3D ELBE Simulator for Educational Use

André Arnold
Forschungszentrum Dresden - Rossendorf e.V.
Dresden Region, Germany
e-mail: a.arnold@fzd.de

Abstract[1]

The paper presents a novel approach to communicate scientific issues based on realistic and interactive 3D simulations. To demonstrate the versatility of this concept, a virtual copy of the electron accelerator ELBE, situated at the Forschungszentrum Dresden Rossendorf (FZD), was created. For programming we used the Open Source Engine Ogre3D, the modeling tool 3Ds Max and the programming environment Visual Studio. The intention of the work is a freely explorable replica of the reality, which is characterized by the interactivity of most of its components and their mouse click retrievable information. The resulting visualization is to our knowledge so far unique in the world and thus exploring completely new ways of imparting knowledge.

1. Introduction

The communication of scientific issues, results and developments to a broad audience is an important task of the public funded research. Positive encouragement from the public is particularly evident when the research underlying scientific principles and relationships are understood. This requires an audience-appealing presentation that allows a playful access to the knowledge but also provides more comprehensive information. An interactive computer simulation is able do both and due to its kind of discovery learning, the world of science becomes touchable and thus less abstract.

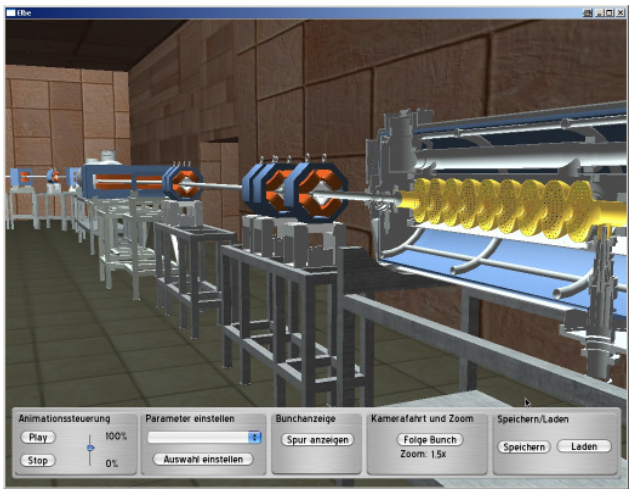
2. Realisation

Computer Graphics and Visualization at the Technical University of Dresden, computer science students and scientific staff of the FZD created such an interactive simulation of the electron accelerator ELBE [1]. Based on the detailed modelling of the main accelerator elements, a copy of ELBE was designed, which represents an abstract but still complete image of the reality. It includes not only magnetic lenses and deflection coils, but also the generation of electron packages in the thermionic electron source, its subsequent compression by means of "Bunchers" and finally the acceleration in the superconducting radio frequency resonators (see Fig. 1).
All the elements are animated and include mouse click retrievable information about its function in text and image. A simplified model of the real physics, on which the individual components of the accelerator are based, offers the possibility to study its effect on the animated electron beam by changing typical parameters such as currents and voltages. It is also possible to load preset simulation parameters.
In particular, a visualization mode can be used in which the electron bunch is followed in a 3rd person perspective on his journey through the entire accelerator. But also the free movement without disturbing the visualization flow is possible at any time. The control is done with the keyboard buttons while the change of perspective can be made via mouse. The desired zoom is selectable by mouse wheel and scene objects can be chosen with the left mouse button. Moreover a user interface is always available to select most of the important functions.

3. Physics behind the Scenes

To ensure a realistic behavior of the application, the well known matrix method for calculating the effect of accelerator beamline elements is implemented in the simulation. This method is based on a linearization of the equations of motion (paraxial approximation), in which both nonlinear effects and the interaction between the particles are not taken into account [2]. The resulting formalism can be easily implemented in the program code but is nevertheless sufficient to describe the physical behavior.
In reality, gauss distributed electron clouds, typically consisting of 480 million electrons, are accelerated. For calculation of this ensemble, a phase-space ellipsoid σ is introduced (Eq. 1). The ellipsoid is 6-dimensional and describes the volume in phase space, which houses all the electrons of one bunch. This can be explained as a "hard" boundary layer, or statistically as a surface of a constant probability of presence (eg. FWHM).

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**



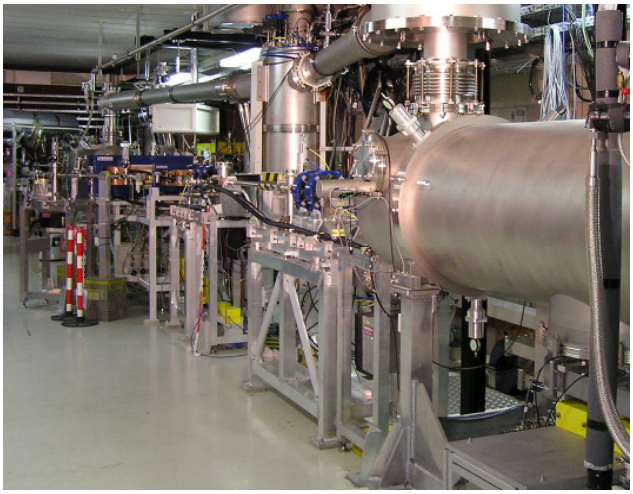


Figure 1. Comparison between the "Ingame" Screenshot of the 3D ELBE Simulator and the Real World

$$\sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & 0 & 0 & 0 & 0 \\ \sigma_{12} & \sigma_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{33} & \sigma_{34} & 0 & 0 \\ 0 & 0 & \sigma_{34} & \sigma_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{55} & \sigma_{56} \\ 0 & 0 & 0 & 0 & \sigma_{56} & \sigma_{66} \end{pmatrix} \quad (1)$$

The elements σ_{11}, σ_{33} und σ_{55} correspond to the square of the three semi-axes, which form the bunch ellipsoid around a central particle (Fig. 2).

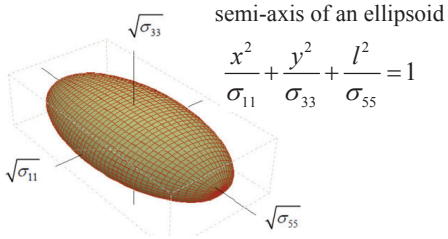


Figure 2. Phasenraum Ellipsoid

The effect of electron-optical elements is now easily computable through the matrix multiplication of the input matrix σ_1 with the transport matrix R of each element (Eq. 2).

$$\sigma_2 = R\sigma_1 R^T \quad (2)$$

The simplest element is a drift space of the length L with its transport matrix R_D and the Lorentz factor γ (Eq. 3). For presenting the transformed bunch after a distance L it is sufficient to extract the three semi-axes of the new matrix σ_2. These values are specified in space with reference to the beam axis and in time relative to a reference particle (central particle).

$$R_D = \begin{pmatrix} 1 & L & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & L & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & L/\gamma^2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (3)$$

For this particle, the beam path has to be determined numerically, in accordance to the equation of motion. In the simulation, this is realized for the motion of the electrons through the high-frequency cavities. The acceleration in this RF field can be calculated by an equation system of two first order differential equations (Eq. 4).

$$\frac{d\beta}{dt} = \frac{q_e}{\gamma^3 m_e c} \underbrace{E_{peak} \sin\left(\frac{2\pi f}{c} z\right)}_{\text{acting electric cavity field}} \underbrace{\sin\left(2\pi f t - \varphi_0\right)}_{\text{time dependence}} \quad (4)$$

$$\frac{dz}{dt} = c\beta \qquad \text{with} \qquad \beta = \frac{v}{c}$$

Where β is the velocity normalized to c, q_e is the charge and m_e the rest mass of an electron. The field strength E_{peak} of the oscillating electric field and its phase relation φ_0 to an arriving bunch can be modified by the user.

To ensure a physically correct, and well observable behavior of the simulation, the speed of light c and hence the frequency f are scaled by a factor of 10^{-9} to 0.3 m/s and 1.3 Hz respectively. The solution of the resulting equation is then realized using the Runge-Kutta method in real-time, which means in parallel to the visualization. The computed relativistic mass increase is shown in Fig. 3. The complete accelerator beamline is finally realized by a sequence of transformation matrices in accordance to the real drift spaces and optical elements.

sich durch ein Gleichungssystem zweier Differentialgleichung erster Ordnung mathematisch beschreiben Eq. 4).

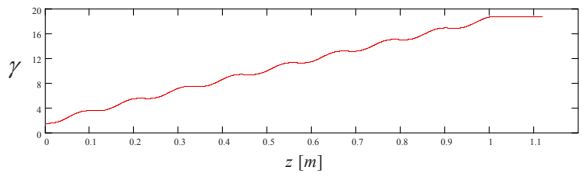


Figure 3. Phasenraum Ellipsoid.

During the program sequence, each bunch is passed from element to element and the resulting ellipse is displayed along the beam path at every frame. The diameter and the length of a bunch, calculated in this manner, are sketched for a beam path starting at the electron source up to the acceleration cavity (see Fig. 4).

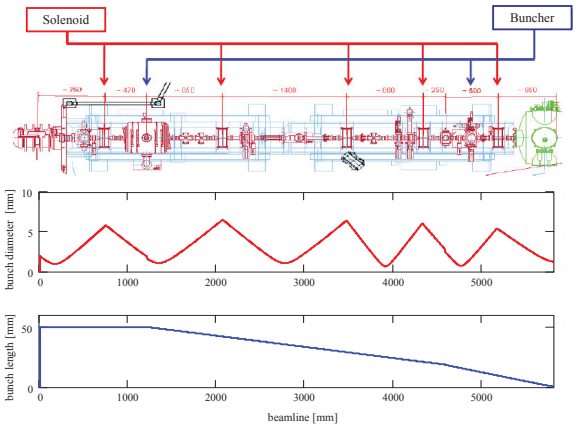


Figure 4. Bunch Diameter (red) and Bunch Length (blue) along the first 5 m of the ELBE Accelerator

4. Program Architecture

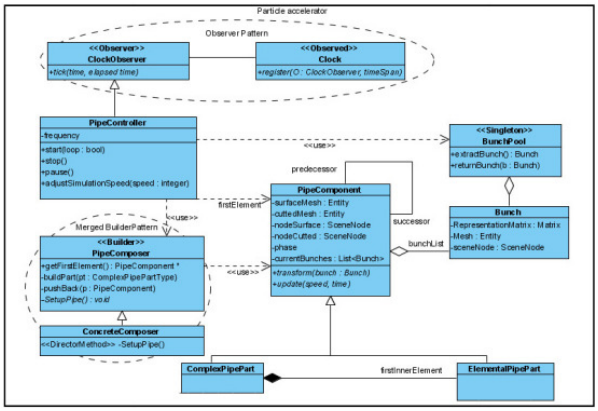


Figure 5. Collaboration Diagram of the Program Architecture

The program design (Fig. 5) essentially corresponds to reality. Thus, each element of the accelerator is described by the base class “PipeComponent” which provides the basic functionality. This class transforms the bunch and passes it to his successor. The “PipeController” works as a steering element because it handles the communication with the GUI, controls the 4. .

simulation process as well as its speed and provides the necessary time synchronization between the elements.
The “BunchPool” contains all free instances of the class “Bunch” and dismisses them as requested by the “PipeController” into the first element. The combination of all individual elements to the final beamline is done by the “PipeComposer”. This class has the ability to generate all the needed instances. The complete element sequence is determined by the director class “ConcreteComposer”.
This program architecture guarantees high flexibility, a rapid construction and a subsequent extension of any beamline section just by changing the setup in the “ConcreteComposer”.

5. Conclusion

In this paper, we firstly presented a new approach for knowledge transfer, based on realistic and interactive 3D computer simulations realized by ELBE simulator. During various public events such as the Open Day at the FZD and the Long Night of Science, the program was presented to the public and also tested by them. The consistent positive feedback from visitors and the interest of the public media (FZD Journal and MDR “Sachsenspiegel”) confirm the suitability of the new approach. In particular the clear presentation of the visualization and its interactivity are emphasized.
Motivated by this success, in 2010 a simulation of the high-performance laser DRACO [3], a separate part of ELBE, was finished. Since both use the same 3D engine and a similar program structure a merger of the two simulations is scheduled. In the future, a project is planned to upgrade the software to be able to visit the entire center for high-power radiation sources of the FZD virtually.

Acknowledgements

The author would like to thank Niels v. Festenberg and his students of the Institute of Computer Graphics and Visualization at the Technical University of Dresden for their long nights of programming and debugging. And a special thank goes to Dr. Jochen Teichert for his support in the physical background.

References

1. P. Michel “The Radiation Source ELBE at the Forschungszentrum Dresden - Rossendorf”. Nuclear Science Symposium. Dresden, 2008, pp.3078–3080.
2. K. Wille: Physik der Teilchenbeschleuniger und Synchrotronstrahlungsquellen. Stuttgart: Teubner-Studienbücher, 1992, pp.75-84.
3. S.D. Kraft, C. Richter, K. Zeil et al: Dose-dependent biological damage of tumor cells by laser-accelerated proton beams. New Journal of Physics, 12 (2010), 085003

On the Influence of Control Flow Properties to Software Error Localization

Andrey Morozov
Institute of Automation
TU Dresden
Dresden, Germany
e-mail: andrew.morozov@gmail.com

Klaus Janschek
Institute of Automation
TU Dresden
Dresden, Germany
e-mail: klaus.janschek@tu-dresden.de

Nafisa Yusupova
Faculty of Computer Science and Robotics
USATU
Ufa, Russian Federation
e-mail: yussupova@ugatu.ac.ru

Abstract[1]

Error localization is one of the most difficult processes in software debugging. Traditionally it is a manual trace-back of the error propagation path from the place where an error was detected up to the error-source. This paper introduces a method for automation of this process. Presented method is based on abstract model of the software, which contains a Markov representation of software control flow, results of the reliability assessment of it independent parts and formal definition of observers. Observers are considered as elements of the software, which report about wrong behavior in the case of error activation or error propagation on them. Combined stochastic analysis of a control flow graph, reliabilities of the software parts and the observation results make it possible to identify the most probable location of the error-source.

1. Introduction

Error localization is one of the most important parts of the software debugging. Developers spend significant fraction of time on it. Traditionally the error localization is a manual search of the error-source. But in recent years much researches has been devoted to automation of this process like [1,2,3]. Our paper presented a new method, based on a control flow analysis. A software error is considered as a logical incorrectness in the some part of the software. It is possible to define six different stages in a life-path of the error:

1. *Error implementation* – a developer introduces the error into the software design or into the code.
2. *Error activation* – some circumstances during the software operation (an unexpected input data, execution mode etc.) leads to wrong behavior of the part of a program, where the error was implemented.
3. *Error propagation* – corrupted data from this part spreads further and leads to the malfunction of some other parts of the software.
4. *Error detection* – somewhere the wrong behavior becomes visible, the error receives the status of a detected error.
5. *Error localization* – maintenance team trace back a propagation path of the error, looking for the error-source.
6. *Error correction* – necessary changes were made in the localized part of the software.

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

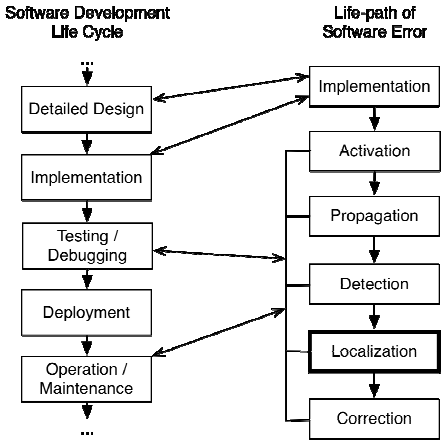


Figure 1. Software Error's Life and SDLC

In the fig. 1 we can see correspondence of the stages of a software error life-path and phases of a *Software Development Life Cycle* (SDLC). This paper addresses the error localization stage.

We introduce a method for probabilistic analysis, which helps to identify the parts of the software application, where the detected error was activated most likely.

The method is based on the Markov model. Software is considered as a system of independent elements. Each element represents an executable part of the code, which can contain implemented errors. The elements are connected according to control flow of the software into the control flow graph. The arcs of the graph are weighted and show probabilities of transitions from one element to another, like in a *Discrete Time Markov Chain* (DTMC).

Some elements of the software under consideration we regard as observers – the software parts where an error will be detected (observed) during the execution, e.g. data checking, output to the log-files or user interface. Using the information from the observers, analysis of the control flow graph and a reliability assessment of the software elements we are making probabilistic forecast of the location of the error-source.

The rest of the paper is organized as follows. Chapter 2 describes the structure of the method. Chapters 3 and 4 discuss the formal definition of the control flow representation. Chapter 5 gives short overview of the models for reliability assessment. Chapter 6 tells about mean number of executions of the element during the execution of the system. Chapter 7 defines the error observation model. Chapter 8 discuss error localization algorithm using an abstract example.

2. Structural overview

Presented method can be well described as a set of several components (fig. 2).

An upper level of the components consists of available data. The error localization stage corresponds to the testing/debugging or operational phase of the SDLC. In these phases both software architecture definition (e.g. UML diagrams) and an implemented code are already existed – this two instances are grouped into the "Software"-block at the fig 2. Another available data – is an error report, created after the error detection. It tells about conditions, which leads to an erroneous situation.

Next level of the structure is presented by a control flow analysis, a reliability assessment and an error report formalization. This is an intermediate level to convert available real instances to the input parameters of an *Error Localization Algorithm* (ELA). The intermediate transformations will be covered only superficially because the goal of the paper is to describe the ELA.

Third level contains input parameters for the ELA. They will be discussed in detail in the next chapters as well as the ELA and it application.

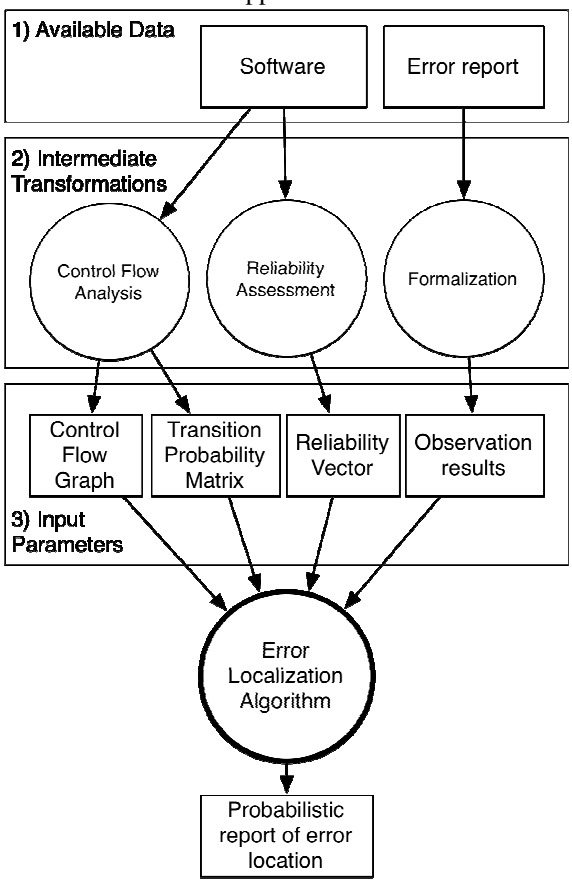


Figure 2. Structure of the Method

The introduced method is a theoretical approach only. It is a start-up for next applied researches. A number of strong assumptions were made to simplify the explanation of the main idea:

1) The execution of the software satisfied to the markov property: probability of execution of the next element depends only on the element, which is executed at the moment.

2) An activated error propagates from the error-source to all elements executed later.

3) Observers always detect an activated or propagated error.

3. Control Flow Graph

A *control flow graph* (CFG) is a mathematical and visual representation of a control flow . As was already mentioned, we consider software as a set of connected executable elements. The existence of the control flow link from element e_i to element e_j means that e_i will be executed immediately after the execution of e_j. CFG nodes represent the elements, arcs - *control flow transitions*.

An existence of more than one outgoing arc denotes a control flow fork after the element's execution. It gives a possibility to model conditional transitions (e.g. 'if' or 'case' statements). Backward arcs denote different types of loops in the software program.

Mathematically we define a control flow graph as follows:

$G^{CF} = [E, L^{CF}]$ - CFG, digraph,

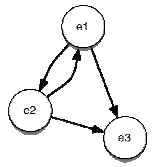
$E = \{e_1, ..., e_{N_e}\}$ - set of CFG nodes,

$L^{CF} = \{\{e_i, e_j\}, ..., \{e_k, e_l\}\}$ - set of CFG arcs.

$\{e_i, e_j\} \in L^{CF}$, if there is a CF transition from element e_i to element e_j.

The example shown in Table 1 represents a CFG of software, which consists of three executable elements.

Table 1: Control Flow Graph

CFG	Mathematical Definition
	$G^{CF} = [E, L^{CF}]$ $E = \{e_1, e_2, e_3\}$ $L^{CF} = \{\{e_1, e_2\}, \{e_1, e_3\}, \{e_2, e_1\}, \{e_2, e_3\}\}$

A structure of a CFG can be obtained in a straightforward way from an UML model at the design phase of the SDLC and specified during the further phases [4,5,6].

4. Transition Probability Matrix

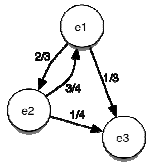
Probability of transition (PT) is a property of a CFG arc. Probabilities of transitions are written in matrix form like a weighted adjacency matrix of a CFG. It is called a *transition probability matrix* (TPM) and denoted by P^{trans}. All elements of the TPM are in the interval from zero to one as well as the sum of the values of each row of the TPM:

$$0 \le P_{i,j}^{trans} \le 1$$

$$0 \le \sum_j P_{i,j}^{trans} \le 1 \quad \forall i, j$$

The TPM is a square matrix of order N_E, where N_E is a number of elements in software under consideration. Value $P_{i,j}^{trans}$ shows the probability of transition through arc $\{e_i, e_j\}$. An example of TPM is shown in the Table 2.

Table 2: Transition Probability Matrix

CFG	*TPM*
	$P^{trans} = \begin{bmatrix} & e_1 & e_2 & e_3 \\ e_1 & 0 & 2/3 & 1/3 \\ e_2 & 3/4 & 0 & 1/4 \\ e_3 & 0 & 0 & 0 \end{bmatrix}$

The probabilities of transitions should be defined for the CFG additionally. As a rule, developers can already predict probabilities of transitions with some precision at the design phase of the SDLC. Technically this information can be added to the UML model via UML stereotyping mechanisms. At further SDCL phases the probabilities of transitions can be concretized using approaches like [7,8].

5. Reliability

Another input parameter for the ELA is a vector of reliabilities of the elements. Reliability is defined as a probability of error activation during the element execution. There are a lot of existing models for reliability assesment, which are aplicable at the different phases of SDLC. According to [9] the models can be devided into the several groups:

- Error seeding
- Failure rate
- Curve fitting
- Reliability growth
- Markov structure
- Time-series
- Nonhomogeneous Poisson process

The question of the model selection is a stand-alone problem and it's not covered in this paper.

Reliabilities for all elements are considered as a known parametrs. Mathematicaly we will denote them by a vector:

$$\overline{R} = [R_1, ..., R_{N_E}]$$

Where, R_i - is the reliability, or the probability of error activation during the execution of e_i.

6. Mean Number of Executions

Before the explanation of the ELA we should define one more property - *mean number of executions* (MNE). It is a property of the CFG's node, which shows an expected number of executions of the element during one execution of the entire software. This property satisfied only for CFGs with the elements, which sum of PTs of outgoing arcs is less than one:

$$\exists e_i \in E, \sum_j P_{i,j}^{trans} \le 1$$

Otherwise it tends to infinity.

MNEs are written as a matrix N. The element $N_{i,j}$ denotes the MNE of e_j, under the condition that the execution was started with element e_i.

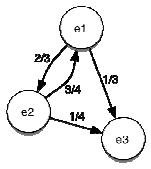
In the terms of Markov model it is the DTMC with absorbing states [10]. In this case MNE matrix can be calculated as a fundamental matrix the chain, as shown in the Equation-1.

$$N = (I - P^{trans})^{-1} \qquad (1)$$

I - is an index matrix in oder N_E.

Table 3 shows the MNE matrix for the previous example.

Table 3: Mean Number of Executions

CFG	MNE
	$N = \begin{bmatrix} & e_1 & e_2 & e_3 \\ e_1 & 2 & 4/3 & 1 \\ e_2 & 3/2 & 2 & 1 \\ e_3 & 0 & 0 & 1 \end{bmatrix}$

A product of element's reliability and the MNE shows *probability of error activation in the element during the execution of the entire software* (PEA). $P_{i,j}^{EA}$ - is the PEA of the element e_j under the condition that the execution was started with element e_i.

PEA calculates as follows:

$$P_{i,j}^{EA} = R_j N_{i,j} \qquad (2)$$

This value will be used as a basic measure for the error localization.

7. Error Observation Model

Some of the element of the software we consider as obsevers. Observer is an element, which reports about wrong behavior in the case of either error activation in the element or error propagation through the element.

Visually we denote observers on the CFG with the loupe sign. Elements e_2, e_7 and e_{10} in fig. 3 are observers.

Each observer after the execution commits a positive or negative note. These notes are available after the error detection and called observation results. One of the possible observation results for the CFG in fig. 3 is shown below:

$$\begin{bmatrix} e_2^+ & e_7^+ & e_2^+ & e_{10}^- & \dots \end{bmatrix}$$

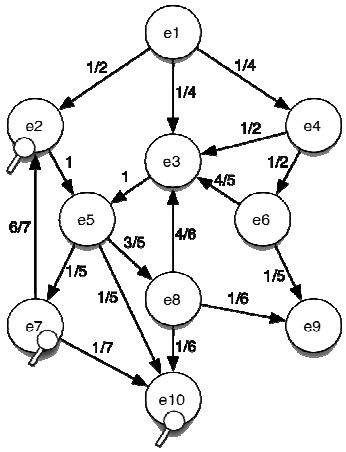


Figure 3. Example of the Control Flow Graph

Observers number 2 and 7 were executed with positive notes. Last positive note was from e_2, and after that e_{10} committed negative note. It means that error was activated in the element which was executed between the executions of e_2 and e_{10}.

8. Error Localization Algorithm

Error localization algorithm will be described with the help of the previous example (fig. 3).

8.1. Problem Definition

Let us suppose, that we have an erroneous situation with the software, which satisfy to the CFG in fig 3. TPM is known as well as the observers - elements e_2, e_7 and e_{10}.

Assume that reliabilities of all elements are the same and equal to r. The observation results are identical with the results from the previous chapter:

$$\begin{bmatrix} e_2^+ & e_7^+ & e_2^+ & e_{10}^- & \dots \end{bmatrix}$$

The problem is to find out which of the elements is the most probable error-source.

8.3. Solution

We can see that the last positive note was committed by element e_2 and the first negative – by element e_{10}, so the error was activated in the element which was executed between them. Also we know that element e_7 was not executed in this interval. So we are interesting only in the part of the execution sequence between elements e_2 and e_{10}.

Therefore we are reducing the control flow graph, to examine only the probable execution scenario between the launches of element e_2 and element e_{10}.

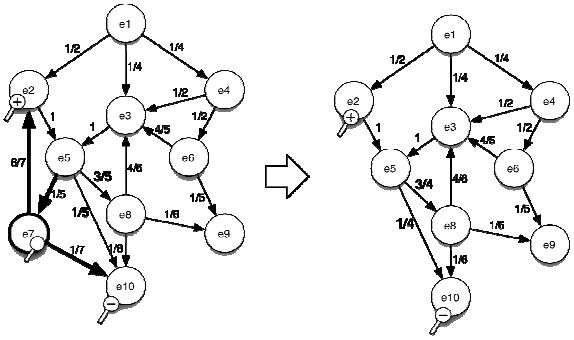


Figure 4. ELA: Step 1

Step 1. Excluding of inactive observers (in our case - only e_7) and their input and output arcs (fig. 4).

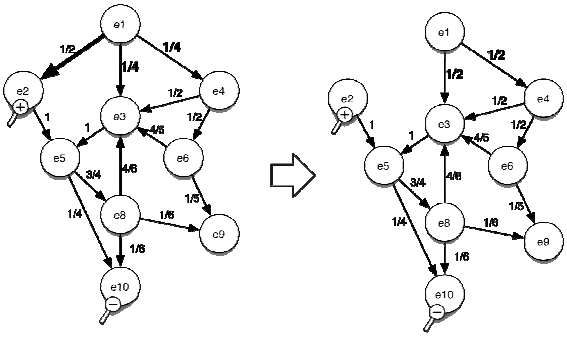


Figure 5. ELA: Step 2

Step 2. Making the last observer with the positive commit (element e_2) an initial element of the CFG through the removal of incoming arcs. Making the first observer with negative commit (element e_{10}) a finite element through the removal of outgoing arcs (fig. 5).

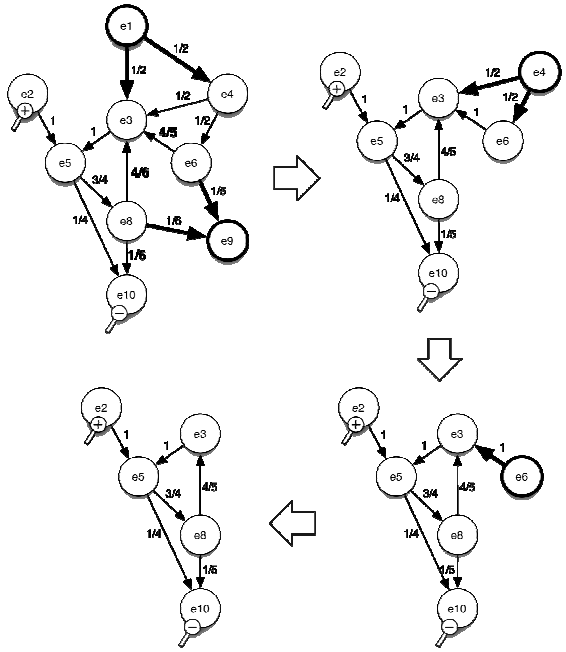


Figure 6. ELA: Step 3

Step 3. Remove all elements, which are not ascendants of e_{10} and descendants of e_2 at the same time. In our example it can be done by recurrently remove of all elements without incoming arcs except e_2 and all elements without outgoing arcs except e_{10}. This operation gives us a reduced CFG. The reduced graph shows all probable execution scenarios, which start with the launch of element e_2 and finish with the launch of element e_{10} (fig. 6).

Step 4. Calculation of the MNE matrix for the reduced CFG with the help of Equation-1.

$$N = \begin{bmatrix} & e_2 & e_3 & e_5 & e_8 & e_{10} \\ e_2 & 1 & 3/2 & 5/2 & 15/8 & 1 \\ e_3 & 0 & 5/2 & 5/2 & 15/8 & 1 \\ e_5 & 0 & 3/2 & 5/2 & 15/8 & 1 \\ e_8 & 0 & 2 & 2 & 5/2 & 1 \\ e_{10} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The first row shows MNEs of the element in the case when the execution begins with the actuation of e_2.

Step 5. Calculation of PEAs using Equation-2, under the condition that software execution starts with e_2.

$$P_{2,i}^{EA} = \begin{bmatrix} r & \frac{3}{2}r & \frac{5}{2}r & \frac{15}{8}r & r \end{bmatrix}$$

Result: MNA matrix shows that in the all-probable execution scenarios, from the execution of e_2 to the e_{10}, element e_5 executed most often. It is the most probable error-source because of the equality of the reliabilities. Element e_5 should be checked by software maintenance team first of all. If it operates correctly, then should be checked elements e_8, e_3 and e_{10} respectively.

9. Conclusion

A theoretical method for software error localization with the help of a control flow analysis was introduced in the paper. The method is based on the Markov model and shows the influence of the control flow into the software reliability aspect. The method is a state-based and applicable to the CFGs with different types of cycles unlike the path-based approaches.
It can be adapted not only for software analysis, but for any type of a system of executable elements with defined control flow. We plan to use it as a part of error propagation analysis of embedded systems of software and hardware elements, which are widely used in the areas of automation, robotics, mechatronics, automotive and aerospace industry.
There are a lot of trade offs in defining the elements, extracting of CFG structures and calculation of the reliabilities and probabilities of transitions. Numerous papers concern these issues, but just few experimental studies, with real results, have been published so far.
Therefore the main direction of the future development of the method is its adaptation for real systems.

References

1. P. Arumuga Nainar, T. Chen, J. Rosin, and B. Liblit. Statistical debugging using compound boolean predicates. In International Symposium on Software Testing and Analysis (ISSTA), pages 5-15, 2007.
2. B. Liblit, M. Naik, A. X. Zheng, A. Aiken, and M. I. Jordan. Scalable statistical bug isolation. In Conference on Programming Language Design and Implementation (PLDI), pages 15–26, 2005.
3. C. Liu, X. Yan, L. Fei, J. Han, and S. P. Midkiff. SOBER: statistical model-based bug localization. In joint meeting of the European Software Engineering Conference and ACM SIGSOFT Symposium on the Foundations of Software Engineering (ESEC/FSE), pages 286–295, 2005
4. Vittorio Cortellessa and Raffaela Mirandola, “Deriving a queueing network based performance model from UML diagrams,” in Proceedings of the 2nd international workshop on Software and performance (Ottawa, Ontario, Canada: ACM, 2000), pp. 58-70.
5. Peter King and Rob Pooley, “Derivation of Petri Net Performance Models from UML Specifications of Communications Software,” Specifications of Communication Software”, Proceedings of XV UK Performance Engineering Workshop, 2000, pp. 262-276.
6. Christoph Lindemann, Axel Thümmler, Alexander Klemm, Marco Lohmann and Oliver P. Waldhorst, “Quantitative system evaluation with DSPNexpress 2000,” in Proceedings of the 2nd international workshop on Software and performance (Ottawa, Ontario, Canada: ACM), 2000, pp. 12-17.
7. Sherif M. Yacoub, Bojan Cukic, and Hany H. Ammar, “Scenario-Based Reliability Analysis of Component-Based Software,” in Proceedings of the 10th International Symposium on Software Reliability Engineering (IEEE Computer Society), 1999, p. 22.
8. Swapna S. Gokhale, W. Eric Wong, Kishor S. Trivedi and J. R Horgan “An Analytical Approach to Architecture-Based Software Reliability Prediction” in Proc. 3rd Int’l Computer Performance and Dependability Symp., 1998, pp. 13-22.
9. Hoang Pham, System Software Reliability, first ed. Chapter 5, (Springer, 2006).
10. Charles M. Grinstead и J. Laurie Snell, Introduction to Probability, Chapter 11, 2nd ed. (American Mathematical Society, 1997).
11. Apel, S., “Software reliability growth prediction - state of the art,” Kaiserslautern, 2005, VII, 38 pp.
12. Katerina Goševa-Popstojanova and Kishor S. Trivedi “Architecture-based approach to reliability assessment of software systems,” Perform. Eval. 45, no. 2-3, 2001, pp.179-204.
13. Arshad Jhumka, Martin Hiller, and Neeraj Suri, “Assessing Inter-Modular Error Propagation In Distributed Software,” symp. On Reliabile Distributed Systems, 2001, pp. 152-161.
14. Martin Hiller, Arshad Jhumka, and Neeraj uri, “An Approach for Analysing the Propagation of Data Errors in Software,” in Proceedings of the 2001 International Conference on Dependable Systems and Networks, 2001, pp. 161-172.
15. B. Littlewood, “Stochastic reliability growth: A model with applications to computer software faults and hardware design faults,” in Proceedings of the 1981 ACM workshop/symposium on Measurement and evaluation of software quality, 1981, pp. 139-152.
16. B. Littlewood, “A reliability model for Markov structured software,” в Proceedings of the international conference on Reliable software , 1975, pp. 204-207.

Education Process Management: a Problem of Curriculum Comparing

Anastasia Markelova
Chair of Computing Mathematics and Cybernetics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: markelova.anastasia@gmail.com

Dr. Ing. Olga Smetanina
Chair of Computing Mathematics and Cybernetics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: smoljushka@mail.ru

Abstract[1]

The article is addressed to the problem of education process management from the point of view of educational programs. It is analyzed foreign educational programs and compared to Russian's programs for retrieving a useful experience. The emerging problem of curriculum's heterogeneity proposed to solve using regular expressions. As result, the example of comparative analysis is given, based on comparing the curriculum of the leading Belgian institute in the area of informatics – Institut Paul Lambin, University of Luxembourg, and the educational program of Ufa State Aviation Technical University (USATU), Russia.

1. Introduction

Growing competition in educational services between universities and Russia's joining to the Bologna process makes the improvement problem of educational process management very actual, particularly in the field of educational programs. In order to attract promising international students as well as Russian students and at the same time to fulfill the obligations stated in Bologna agreement [1], it is necessary to develop competitive educational programs and represent them on the market in a right way.

The analysis of foreign curricula advantages and their following implementation to Russian's educational programs as a next step lets to get standardized curriculums, which keep the Bologna terms and respond to the third Russian state education standard at the same time [2].

As an example of international European experience, it were taken Institute Paul Lambin (Belgium) and University of Luxembourg for this article. This leading Belgian Institute in the sphere of informatics and young, but famous for its quality of studying yet, University of Luxembourg, are comparing here with Ufa State Technical University (Russia).

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

1 The improvement problem of educational process management

It can be studied from the different points of view. There are many directions in educational process, which are possible to be improved. On the figure 1 represented a schema with potential points for optimizing.

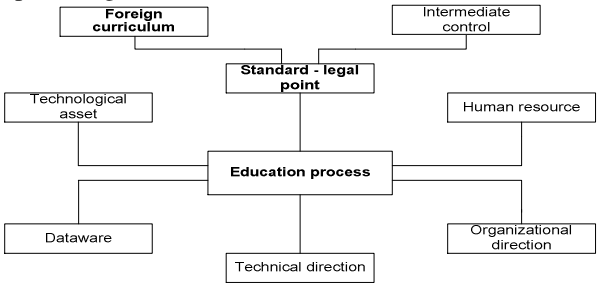


Figure 1. Schema with potential points for optimizing

In this article is discussed a standard- legal point, which means a set of legislative acts from the law about education and state educational standards.

An educational program is the first point to optimize and improve, because exactly from this point a student begins to deal with from his very first day at the university. If the program is not responding to the actual reality, for example, there are no necessary subjects, or the volume of subject is not enough, then there is a strong probability to get in the future a specialist not demanded on the market and incapable for mobility[3].

A curriculum needs to be improved according to Russian state education standard of the 3rd generation. The aims of this standard are an implementation of the Bologna terms in the universities: an education system with 3 degrees- bachelor, master and doctorate, ECTS (European Credit Transfer System). And one of the most important points is to give more freedom to students in choosing the disciplines. Especially here we need the experience of the European countries.

There is also one subpoint in the standard- legal point that is not going to be discussed here- an Intermediate control. The definition is that a test, estimating the

knowledge of the students on some level of the education. It means that we evaluate the knowledge on basis of non-optimized educational programs, which are not interesting for a current task. We need to remake a curricula first, and then estimating their adoption.

Thereby, there are many factors, which influence on a new educational plan formation. All these factors are presented on the figure 2.

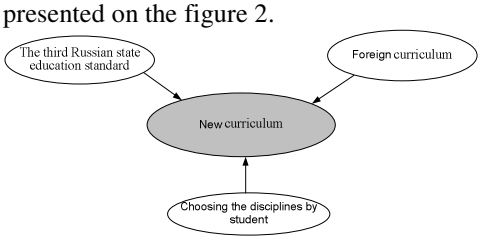


Figure 2 Factors influencing on a new educational plan formation

2 A comparative analysis of educational programs of Russian and foreign universities on the basis of the regular expressions.

One of the main tasks in comparative analysis of curriculum is processing of symbol information. For example, the name of the discipline can consist of the set of the words and digits, and the load can consist only of digits. For the following creation of data base, the processing of such information is enough labor-intensive task. Regular expressions were used for the solution of this problem. It is actual retrieval system of text wraps in the electronic document, based on special pattern remembered system for searching. The pattern giving the rule for searching is named "mask" sometimes [4].

The regular expressions are used for compact description of some aggregates with the help of the patterns, without the necessity of enumeration of all of the elements of this aggregate. It is used the special syntax for making the patterns. The examples of special symbols and operators are in the Table 1 [5].

So the regular expressions can be applied for:

а) searching of substring, satisfying to the pattern of regular expression, in the line.

б) searching and replacement of substring, satisfying to pattern of regular expression, in the line.

в) check-up for equivalence the defined line to the pattern.

г) retrieving ,substring satisfying to the pattern of the regular expressions , from the line.

Table 1. Examples of special symbols and operators

Special symbols:		Operators	
^	the beginning of the line	{n}	Repeat n time
\b	the beginning of the word	|	Previous or next pattern(logical "OR")

The examples of the regular expressions, applying at the analysis of educational programs:

«(\w|\s|\d|'|/|-|\(|\))+($|\n|\r|\t|;|,|.)» - for getting certain key words from the set of the key words of the discipline; «[^\r|\n|\a]» - for getting the name of the discipline without the symbols of formatting.

2.1 The comparison results

In this article are compared the educational program of Paul Lambin's Institut [6], which is one of leading universities in Belgium in the field of informatics, the educational plan of University of Luxembourg [7]and the educational program of USATU(Ufa State Aviation Technical University), one of significant universities in Russia in the same field.

The comparison analysis is done for all blocks of educational programs: humanities and social- economic disciplines (HaSaED), mathematical and natural sciences disciplines (MaNS), general professional disciplines (GPD) and special disciplines (SD). Table 2 reflects the comparison results by the blocks of disciplines.

Conclusions: The comparison by blocks of disciplines shows that comparing with USATU in Belgian and Luxembourgish universities is given 2 times less for

Table 2 Comparison results by the blocks of disciplines

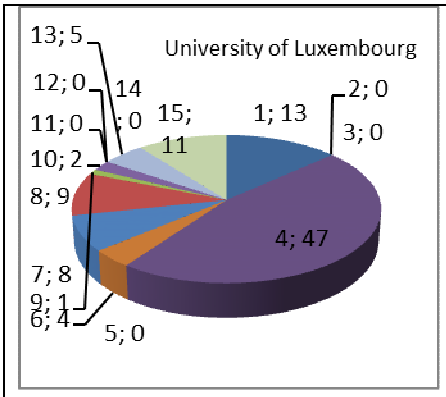
Block	Paul Lambin's Institut, Belgium (%/ECTS)	University of Luxembourg (%/ECTS)	USATU, Russia (%/ECTS)	Conclusions
HaSa ED	10%/18	11%19	26%/ 50	For studying of humanities and social-economic disciplines is given 2 times more in Russia
MaNS	13%/24	13%24	32%/61	For studying the mathematical and natural sciences is given 3 time less in Belgian Institut and at the University of Luxembourg
GPD	45%/81	37%66	28%/54	Learning of general professional disciplines takes almost have of foreign programs, in contrast to Russia, where GPD block takes one third.
SD	32%/59	39%71	14%/26	The special disciplines are learned 2 times less in USATU then in Belgium and Luxembourg

humanities and social- economic disciplines.
Mathematical and natural sciences disciplines in Belgium and Luxembourg are studied 3 times less form the total of course- time. The block of general professional disciplines are very different form Russian program, because in European plans it takes 2 times more than USATU.
But for special disciplines, contrary, 2 times more than Russian university.
The next step is to compare the fields of knowledge.

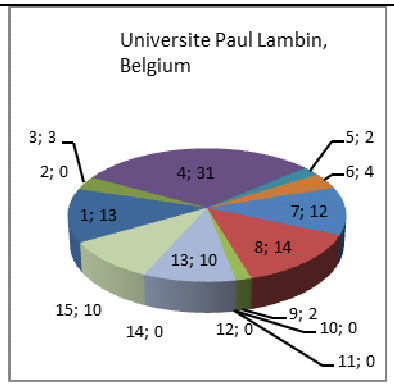
There are 15 fields, where all subjects are grouped by area, for example, mathematics, or computer vision. All humanities disciplines are combined in one area.
Table 3 shows the comparison results by the fields of knowledge. The results of comparison by the fields of knowledge are represented on the diagrams (the image 4.1, 4.2, 4.3). The first digit means the area of knowledge, and the second- its part in percentage form the total amount of the disciplines.

Table 3 Comparison results by fields of knowledge

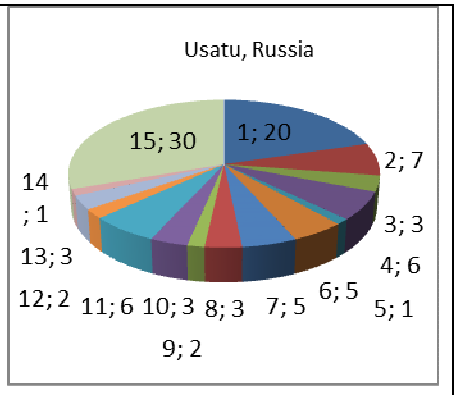
Field	Paul Lambin's Institut, Belgium (%/ECTS)	University of Luxembourg (%/ECTS)	USATU, Russia (%/ECTS)	Conclusions
Mathematics	13%/24	13%/24	20%/39	Mathematics is learned more in USATU
Physics	0%/0	0%/0	7%/14	In Belgium and Luxembourgish universities physics are not studied
Computer vision	3%/6	0%/0	3%/7	Computer vision are studied the same time in USATU and Paul Lambin's universities, but not studied at all at Luxembourgish
Software design	31%/56	47%/84	6%/12	For software design is given 5 times more in Belgium and almost 10 times more in Luxembourg, than in Russia
Artificial intelligent	2%/4	0%/0	1%/2	Artificial intelligent is not studied at all Luxembourg university and studied a bit in other 2 universities
Hardware design	4%/7	4%/7	5%/9	Hardware design is studied in every researched University equal time
Modeling	12%/21	8%/14	5%/10	It is given more time for modelling in Paul Lambin's university
Computer network	14%/25	9%/17	3%/7	Computer networks are studied 5 times more in Belgian university and 3 times more in Luxembourgish university
Information security	2%/3	1%/2	2%/3	Information security is studied equally
Electrical engineering	0%/0	2%/4	3%/7	There is no electrical engineeting course in Belgium university. But it is studied in USATU and in Luxembourg university the same time
Control theory	0%/0	0%/0	6%/12	Control theory is not studied in foreign universities
Reliability	0%/0	0%/0	2%/4	Reliability is not studied in foreign universities
Operating systems	10%/18	5%/9	3%/6	It is given 3 times less in USATU than in Belgian university
System software	0%/0	0%/0	1%/3	System software is not studied in foreign universities
Humanties disciplines	10%/18	11%/19	30%/58	In Belgian and Luxembourgish universities there are 3 times less



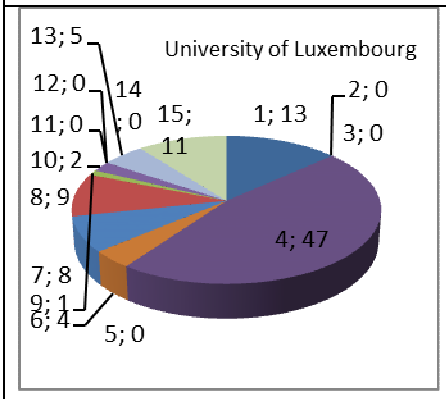
The image 4.1. The field of knowledge of the educational program of Paul Lambin's Institut



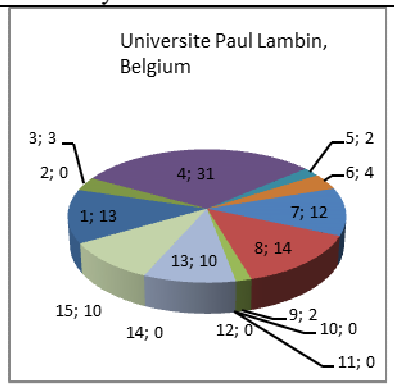
The image 4.2. The field of knowledge of the educational program of the Luxembourg university



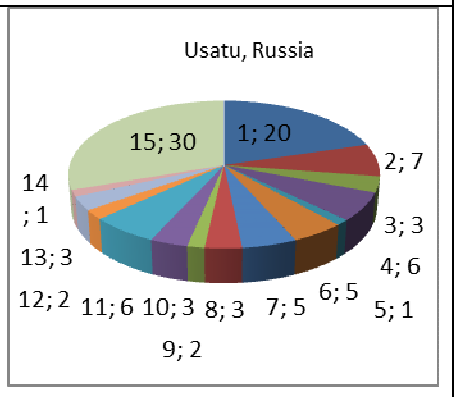
The image 4.3. The field of knowledge of the educational program of USATU



The image 4.1. The field of knowledge of the educational program of Paul Lambin's Institut



The image 4.2. The field of knowledge of the educational program of the Luxembourg university



The image 4.3. The field of knowledge of the educational program of USATU

The field of knowledge on the images: 1 – mathematics, 2 – physics, 3 – computer vision, 4 – software design, 5 – artificial intelligence, 6 – hardware design, 7 – modeling, 8 – computer networks, 9 –information security, 10 – electrical engineering, 11 – control theory, 12 – reliability, 13 – operating systems, 14 – system software, 15 – humanities disciplines.

The comparison by fields of knowledge shows that in USATU mathematics studied deeper. General professional disciplines (GPD) and special disciplines (SD), such as software design, computer network, are studied few times more in Belgium and Luxembourg. But comparing to USATU, there are no some disciplines in studying plan in both foreign universities: physics, control theory, reliability, system software. Also it is given three times less for humanties disciplines there.

3. Conclusions

The perfection of educational process management is a key moment for improving the quality of education and making the competitive educational programs,

therefore it will attract foreign students. It is impossible to do without educational program optimization. It should be done according to third Russian state education standard, obligations stated in Bologna agreement and of course we should take all positive sides, gotten from the comparative analysis of Russian and foreign educational programs.

As an example of international European experience in this paper was taken the educational programs analysis of Institute Paul Lambin Belgium) and University of Luxembourg. From the comparison it is possible to make the conclusion that the educational programs of foreign universities are quite different from the program of USATU. As assumption it can be suggested to reduce till minimum the block of humanities and social- economic disciplines (HaSaED) and due to that to increase the preparation in general professional disciplines (GPD) or special disciplines (SD).

References

1. http://www.ond.vlaanderen.be/hogeronderwijs/bologna
2. http://www.edu.ru/db/portal/spe/3v.htm
3. http://technical.bmstu.ru/umo/index.php?rzd=15&rzdid=15
4. Aho, Alfred V. "Algorithms for finding patterns in strings". The MIT Press, 1990.
5.
6. Friedl, Jeffrey. "Mastering Regular Expressions". O'Reilly, 2002.
7. http://www.ipl.be/
8. http://wwwfr.uni.lu/fstc

Alumni Management Systems and Supporting Software

Liliya Iskhakova
Department of Administration
FZ Dresden-Rossendorf
Dresden, Germany
e-mail: l.iskhakova@fzd.de

Nafisa Yusupova
Faculty of Computer Science and Robotics
USATU
Ufa, Russian Federation
e-mail:yussupova@ugatu.ac.ru

Bjoern Wolf
Department of Support Staff
FZ Dresden-Rossendorf
Dresden, Germany
e-mail: b.wolf@fzd.de

Abstract[1]

Today's research business is characterized by cuts in government funding and intense international competition, which lead to an increasing pressure for research institutions to find new ways of generating income. Hence, research commercialization, the development of linkages between companies and research institutions and getting additional financial support for them has become a major issue. However, commercialization practice and the existence of linkages are still insubstantial.

Comparing fundraising practices of USA and UK with the experiences in this field in other European countries (EC) and Russia the authors conclude that Alumni Management System (AMS) could be used to solve some important problems of technology transfer in each country and information technology plays significant role in building a strong relationship with alumni.

Keywords: fundraising, technology transfer (TT), Student Relationship Management (SRM), Alumni Management Systems (AMS), alumni relationship (AR), information technology (IT), Alumni Association (AA), key success factors (KSF).

1. Introduction

The starting point for this study rests on the understanding connected with uncomfortable truths. The first truth is that in knowledge-based societies research is increasingly becoming the basic determinant of growth and prosperity. The second truth is that Europe today seriously under-invests in research. According to an investigation led by an expert group in 2005, 1.9% of gross domestic product (GDP) was spent for research and development (R&D) in the European Union (EU) as opposed to 2.6% in the US and 3.2% in Japan. This gap corresponds to € 120 billion a year when comparing the EU with the USA, 80% of which is explained by the difference in business spending in R&D [1, p.20].

Today's market has a pressure and a rapid change of competition. To be successful in the industry companies must develop linkages with other firms and research institutions. As a result of increased national and international competition and cuts in government financial support, research institutions are urged more and more to find new ways to generate income. Third, governments have an interest in promoting successful relationships between research institutions and industries. Based on these market conditions, literature about technology transfer, innovation and R&D management multiplies rapidly [2, p.2].

Technology transfer (TT) is the practice of transferring scientific findings (research and fundamental discoveries and so on) and knowledge about market needs from one organization to another one for further development [3].

There are different ways by which universities and research centers could collaborate with companies for the TT.

Great percentage of US and UK university budgets come not from tuitions or state funding but from fundraising (philanthropic sector) especially from alumni [4, p.3]. For the TT it is more important to get money from alumni who work in the industry. Based on the preceding knowledge this paper makes a thesis that AMS will help to solve some important problems of technology transfer and all universities and research centers should develop strong relationships with their alumni.

The paper is organized as follows: Section 2 describes the concepts of "Fundraising" and "Alumni management", and explains factors that have an influence on implementing AMS; Section 3 describes the significant role of the IT for AMS, and compares

[1] **Proceedings of the International Workshop "Innovation Information Technologies - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

different Alumni software; Section 4 concludes the paper.

2. 1 Fundraising

There is a number of alternative ways for universities to relate to prospective donors that can be specified in four different "models" of interaction such as the Major Gift Model, the Foundation Research Model; the Multi-mode Model; the Alumni Model [1].
They are distinguished in terms of issues such as donor types, the university actors taking the lead in fundraising, the degree to which specific donors are targeted, the extent to which donors specify the use of donations and so on. Fig. 1 locates these four "modes" in a diagram defined by the donor types along X-axis and the lead university actors in fundraising along a Y-axis. The degree to which donors establish the use of their donations is increasing upwards and to the right along a Y-axis.
At one end of the spectrum, the "Major Gift" model focuses on the efforts made to attract donations from extremely wealthy individuals. The "Foundation Research" model describes everyday activity of researchers seeking financial support. Typically, researchers make applications for grants, PhD programs or scholarships from large research foundations.
The "Multi-mode" model describes a variety of different types of interaction between universities and smaller research foundations.

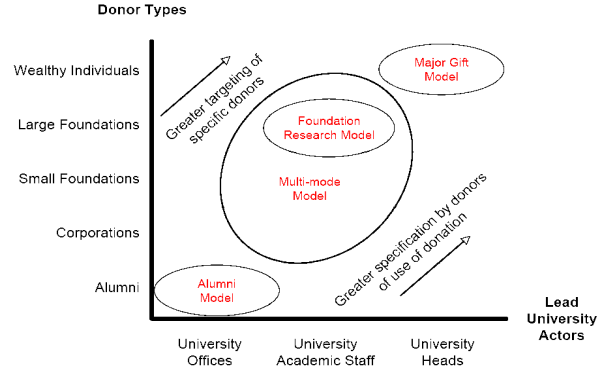


Figure 1. The models of interactions between universities and private donors [1, p. 63]

At the other end of the spectrum, the "Alumni" model refers to university alumni (former student). The discussion about all models would exceed the limit of this paper. Therefore, only the last model will be described in the following section.
Fundraising in US universities compared with other countries has the highest level development. American philanthropy has a very long history therefore people are willing to give. According to the Foundation Center statistics there are 88,000 philanthropic foundations of different sizes in USA. In 2003 $30.0 billion have been donated in USA where 24.5% went to education [4]. In 2005 contributions to US universities from individuals (including rich people and alumni) reached 76.5%, foundations – 11.5%, corporations – 5.3%, bequests – 6.7%. Total benefit from this fundraising was $260.28 Billion [4, p.3] (Fig.2).

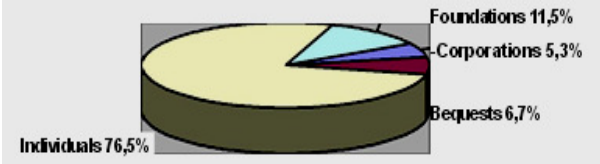


Figure 2. The Foundation Centre Statistics

As we can see from this picture, the amount of individual donors is the overwhelming majority, that's why development offices focus mostly on individuals, especially on alumni.
Philanthropy is a potentially important source, but it is not nearly as well developed in Europe and Russia as in the US.
According to a sociological study of the Russia Centre for public opinion survey which was conducted in 2006 in Russia, 40% of the respondents showed their potential interest in charity and philanthropy, 28% mentioned that they were willing to do something themselves already now, 12% said they could participate only if philanthropy will be supported and highly recognized by government and society. 41% think that philanthropy is just for rich people. Those figures are much more optimistic than in 2001, when just 25% showed potential interest in philanthropy and said that they would be ready to give $ 30 per year [3, p.3]. These results show that although fundraising is a challenge for Russian nonprofit organization, the situation is not hopeless.
Comparing with Russia, the EC fundraising works better but worse than in the US. This lower giving in Europe is clearly related to the European institutional context, with its high taxes and its tradition of public spending for education and research. Although there is a culture of giving in Europe, it is not generally so for education or research [1, p.8].
Despite all different development of fundraising in Russia and EC the general principles and methods of work with individuals, companies and foundations developed in US are the same and many of them can be applied to any advanced country.

2.2 Alumni Management

In 2008 JISC Organizational Support Committee (JOS) has investigated the ways in which institutions manage their relationships with students and alumni. JOS has considered SRM and AMS as a union. This union has been regarded as the intelligent handling of communications between an institution and its students. Within the field of Business innovation, JOS described the SRM and AMS as the development of strategies, policies, and use of IT to support institutions, build and manage relationships with students and alumni through a range of

interactions and engagements during the lifecycle [5, p.4,5].

To create AMS it's necessary for an organization to set up an Alumni Association (AA). AA is to support the university or the research centre by fostering the spirit of loyalty among its alumni, to provide communication links between alumni and the university or the research centre, to increase the sense of pride alumni have in their university or the research centre and so on [4, p.11].

Let's see what kind of benefits a university or a research center can get from alumni.

Alumni can make the following contributions: donate to their university or research center; be active in alumni events and local chapters; use and purchase alumni services; recommend their university or research center to prospective students; hire their institution's students and graduates; participate in the educational process (mentor, instructor, industry contact); volunteer at events and more [6, p.7]. But for doing all these things alumni have to be interested. Which brings up the important questions: *How does an institution change alumni indifference to interest, and interested alumni into donors? What are the key success factors for implementing Alumni Management?*

There are different ways, which were offered by expert groups and scientists and others, for finding a solution for this problem.

In 2009 a famous communications agency called Precedent [7] created "The engagement ladder" (Figure 4) to show how a university should build relationships with their alumni during lifetime. The five stages (from strangers to a life partners) represent relationships, based on mutual interest and affinity.

Experts [7] said that a student's emotional, financial and professional needs changed as they mature.

Similarly, during the lifecycle students make a different degree of contributions consisting of three areas - giving money, providing time and support, and achieving personal development (Fig.3).

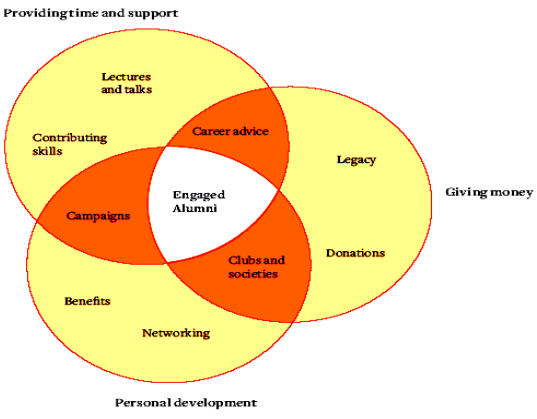


Figure 3. Opportunities for engagement [7, p.6]

There are two example experiences in the AMS field. Fig.4 outlines each of them. The first (1) shows the "traditional" approach of alumni engaging at the last stage. In this period individuals focus on the next stage in their lives and aren't so interested about the life of their university. If they haven't known about the concepts of "giving back" and "alumni" before graduation it would be very difficult for AA to explain about these issues. But if a university consider undergraduates as alumni from the first day (approach 2), former students will be able to understand the concept of "giving back" very quickly and in a more tangible sense.

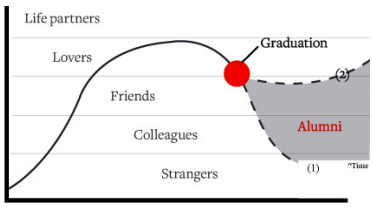


Figure 4. Relationship post graduation [7, p.20]

In 2009 another organization called "CASE's International Benchmarking Survey of Alumni Relations Directors" created a survey and got feedbacks from 90 respondents [8]. These institutions belong to 19 countries, with two-thirds of respondents from the UK, another 12% in continental Europe, and 14% from the rest of the world, including Australia and China.

The survey lets respondents self-assess whether their AR initiative as a "startup" (i.e. in the first year or two of existence and still trying out ventures for the first time), or in an "intermediate" stage (in operation for a few years and still growing, but with a relatively steady portfolio of offerings) or "mature" (those around for several years, with a stable level of staff and a repeated portfolio of offerings). 21% classified themselves as "startups", 57% as "intermediate" and 22% as "mature".

The surveyed groups decided on three measures of success: the numbers of attendees, volunteers, and donors. Having created a list of 28 parameters the experts [10] were searching for correlation (strength of relationship) between these parameters and the three success measures using regression analysis [9]. As parameters were used only things that can be easily measured – number of emails, events, attendees, etc. Intangible factors like quality, satisfaction, and personal experience and so on were not included.

The study resulted in the following main concepts: if you want to increase attendees, then have sufficient staff and budget to put on more events and reunions, rather than increase the size of the same existing events; if you want to increase the number of your volunteers, then be able to connect with them via email or e-newsletters and have a frequent magazine; and if you want to increase the number of donors, then do not expect your database, programs, resources, or events to provide the magic bullet, but be prepared for a long haul, as years of asking and years of having a magazine

appear to be the only variables significantly correlated with donors [8, p.3].
Having studied the result of the survey of 2007 and 2008 the experts [10] found out that if all variables were compared in the same year then the lack of significant correlations between AR variables and donors would be discovered. But if all variables are compared over time, then statistically significant relationships will be found: the number of attendees in 2007 is strongly correlated with the number of donors in 2008, and having donors in 2007 is correlated with having them in a year later (table 1). The message here: AR operations can pay off in terms of donors but over time, not immediately.

Table 1: Correlations between the measures of success themselves, over time

		2007 Attendees	2007 Volunteers	2007 Donors
2008	r	0.326	**0.635****	**0.340***
Attendees	N	35	27	37
2008	r	-0.048	**0.421***	-0.-41
Volunteers	N	35	27	37
2008	r	**0.661****	-0.122	**0.709****
Donors	N	35	27	37

Evidently, universities and research centers are complex unique organizations, making it inappropriate to provide 'of-the-peg' solutions. For example, to create AMS for research centers like Forschungszentrum Dresden-Rossendorf (FZD, Dresden, Germany) it ts necessary to keep in mind the following facts: AMS should keep in contact with former employers, workers or students who conducted their internships or wrote their diplomas there. Therefore, in the USA and UK almost each university has its own alumni program [12,13].
In spite of a lot of differences, SKF might be similar for all education and research organizations.

3. ICT technology for Alumni management system

As explained in Section 2 we could understand that IT plays an important role to create strong AR. For Alumni association it would be much easier to maintain a communication with their former students and to do such work as sending newsletters, magazines, keeping and updating alumni data and informing about events and so on if they would use a database.
But as alumni management software abound in the market, ranged from the basic tools to ones that offer all sorts of additional features and functionality, it is very difficult for universities or research centers to make a decision on whether to buy (off-the-shelf system) or build(tailor-made system)[13].
Many of the larger charities have built their own systems, often at a considerable expense of time, and have them supported by in-house ICT [14]. Leading 'off the shelf' products such as Raiser's Edge [15] and ThankQ [16] can provide some, but not all, of the information management requirements. To find out about the most popular systems in the UK in 2008 Alan Paull described the result of his survey based on the answers from 40 UK institutions [5, p.19-20]. The survey asked users' perceptions of the system and how well the systems in use are meeting business needs. Also of interest was the ease with which data could be transferred between systems and again users were requested to rate their systems at each stage of the student lifecycle.
The factors 'usability' and 'satisfaction' asked for a rating to be provided on how easy the software is to use, and how satisfied the institution. They were evaluated on the scale beginning with the mark "1" corresponding to "bad" up to the mark "5" corresponding "excellent".
The following result showed that the most popular used product was "Raiser's Edge". About 17 universities have been using it over 9 years. Although two users gave ratings of 5 for both usability and satisfaction, the most commonly allocated rating was 3 for both. A similar pattern was seen for SITS users but this software was used by only 4 universities and about 5 universities built and used a tailor-made system designed in-house.

3.1 Alumni software in Germany

In order to know the opinion of customers about different alumni software some universities and research centers which were members of an interactive network tool called CASE-ListServ (ALUMNI-L) or had taken part in the annual alumni-clubs.net conference in Berlin in 2010 have been surveyed by phone and e-mail. The overwhelming majority used Raiser's Edge [14], ThankQ [15] and Banner [16]. It was found that IntrAlumni and JustSoftware were more popular in Germany.
In order to find out more about Alumni software in Germany the companies IntraWorld and Justsoftware were interviewed by telephone. Based on the information obtained some comparison was developed using the unit quantity.
IntrAlumni [17]**:** In Europe and especially in German Speaking Europe **IntrAlumni** is the leading provider of internet-based alumni management and community and has more than 100 customers.
This software is multilingual and provides a basic community platform for content management, relationship management, administration, networking and knowledge exchanges among the institution and its members. It contains currently nine modules that provide extensions to the core platform (e.g. Events, Fundraising, Payments, Career, Market Place, etc.). The customers receive regular updates (extensions, improvements) twice a year.
The cost of IntrAlumni consists of a one-time implementation fee and an annual operation fee [Fig.6]. The first part depends on the customizing module set, while the second part depends on the number of processed profiles:
Fig. 6 shows the dependence of the total cost of a project IntrAlumni on the number of profiles in the database and project duration (in years).

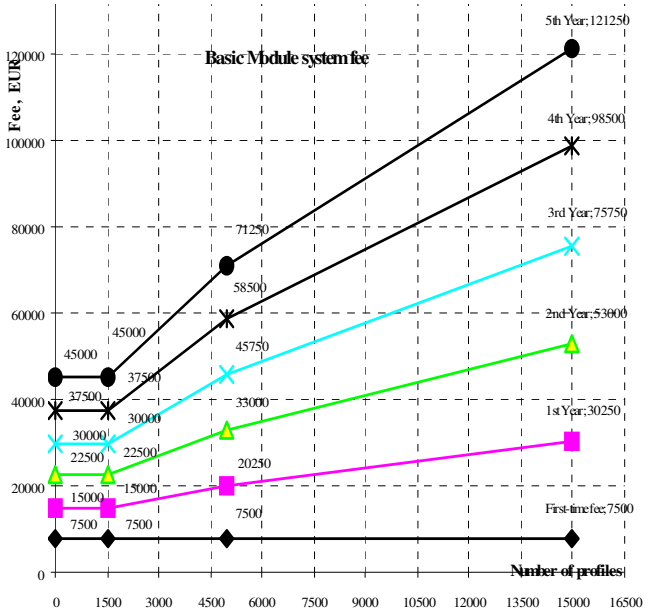


Figure 6. Basic module system fee for five year

The X-axis denotes the number of profiles in the database, which has the following thresholds: 0, 1500, 5000, 15000. When the number of profiles is more than 15000, the fees depend on the individual project and therefore they are not displayed in the graph. The total cost of the project for each year is given by the Y-axis. We assume that during the whole period the number of profiles in the database does not change. The bottom line shows the one-time cost in the implementation of the project (project duration is assumed as 0), the next one above plots the cost of the project with a duration of one year, and so on. The top curve shows the cost of the project for the duration of five years.

To collate the cost of different systems it can also be convenient to compare the partial costs of the systems evaluated for one profile in the database. Partial cost of the profile can be found as the ratio of the whole system cost to the profile number:

$$f_{prof} = \frac{F}{N},$$

where f_{prof} – partial cost of the system, evaluated for one profile, F – whole system cost, N– profile number. Assuming that number of profiles in the system remain unchanged, we obtain the following diagram (Fig.7):

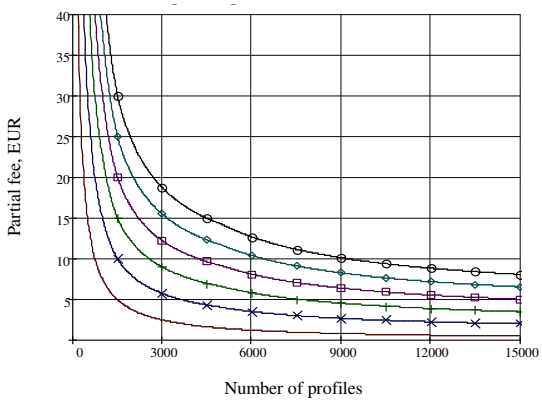


Figure 7. Partial costs for five year (IntrAlumni)

Here, again, each curve corresponds to a certain duration of the project: the lowest corresponds to one year, the highest one to- five years. The chart shows that the more profiles (users), the lower the unit cost of the project, i.e. the higher its efficiency.

JustSoftware [18] has face book-usability combined with key features for Alumni-Management and virtualization of a university (networking structure like folder system of Microsoft allows a virtualized copy of the real university with faculties or different studies) JustSoftware has more than 10 customers. The average product price is 500 euro per month which includes all services. The cost of JustSoftware was analyzed in the same way like the price of IntrAlumni (Fig.8,9).

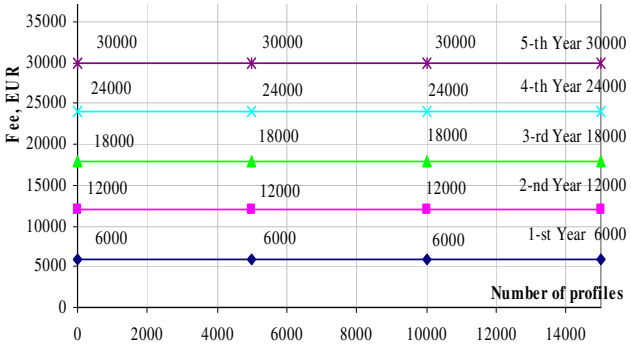


Figure 8. Basic module system fee for five year

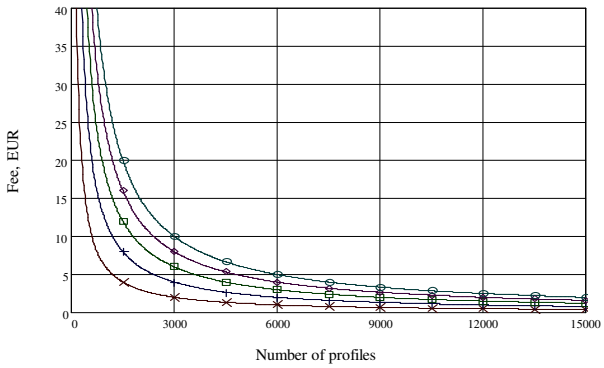


Figure 9. Partial costs for five year (JustSoftware)

To compare the partial fee of these systems we can plot their difference (Fig.10). This graph shows whether one system is more expensive than the other or not. If the curve became negative, the JustSoftware would be more expensive in this region.

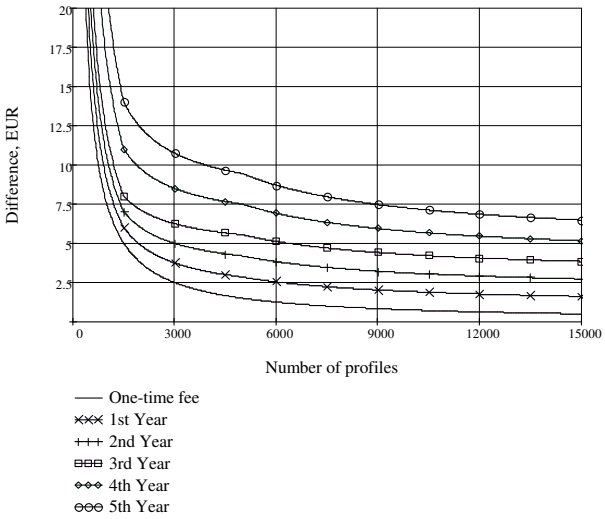


Figure 10. Difference between partial fees JustSoftware and IntrAlumni

From that data it can be deduced that the IntrAlumni system more expensive under all conditions (especially considering additional modules) than the JustSoftware system (the curves remain positive on the whole range from 0 to 15000 profiles and from 1 to 5 years).

4. Conclusion

The paper contains a number of recommendations, the best practices in the USA, the UK and others countries in the philanthropic sector. This study considers key success factors for implementing Alumni Management aimed at universities and research centers to raise the level of fundraising for research within Europe and Russia. The paper shows the prominent role of IT in this field. Famous alumni software in Europe (in more details in Germany), the USA, the UK are described and compared. Future work includes the development of a survey in order to find out the new KSF for AMS, good methodology to compare different modern Alumni software (software created by SAP will be considered). Future study will create expert systems and will provide that AMS will help to solve some main problem of the technology transfer. In the end AMS will be created for the Forschungszentrum Dresden-Rossendorf (FZD, Dresden, Germany) and the Ufa State Aviation Technical University (USATU, Ufa, Russia).

According to [9, p.24] AMS will not only be benefit for the alumni development team, but will also contribute to the objectives of many other departments of the university and the research center.

References

1. http://ec.europa.eu/invest-in-research/pdf/download_en/final_report_v2_en.pdf
2. Key Success Factors for Research Institutionsn Research Commercialization and Industry Linkages: Outcomes of a German/Australian cooperative project T. Baaken ,C. Plewa
3. http://www.betterworldproject.org/tech_transfer.cfm
4. Karpova St.Petersburg State University "Case Study: Alumni Relations, Fundraising and Development in US Universities", IREX University Administration Support Program – 2006
5. A. Paull, D. Chambers "JISC landscape study relationship management"-2008
6. Alumni Engagement: Research in Action, Ipsos & SAIT Polytechnic Alumni Research: A Case study CASE VIII - Portland, Oregon (17-19 February, 2010)
7. http://www.precedent.co.uk/
8. A summary of the seminar given by Chris Thompson on 19.3.09 " What does successful Alumni Relations "look like"?" CASE AR benchmarking survey.
9. http://en.wikipedia.org/wiki/Regression_analysis
10. British Council USA "Background on top USA HEI Alumni programs", 2009
11. British Council USA «US HEI ALUMNI NETWORKS An Overview",2009
12. Choosing a Database "A guide to help small to medium sized organisations find their way through the process of choosing a database", www.ITsorted.org.uk
13. http://www.preponderate.co.uk/databasedilemmas.htm
14. blackbaud.com/products/fundraising/
15. www.esit.co.uk
16. http://www.sungardhe.com/?LangType=1033
17. http://www.intraworlds.com/
18. www.justsoftwareag.com

The Semantic Analysis of the Documents Data on the Example of the Curricula Comparative Analysis

M.B. Guzairov
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: guzairov@rb.ru

N.I. Yussupova
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: yussupova@ugatu.ac.ru

M.M. Gayanova
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: maya_m@mail.rb.ru

V.A. Kozyreva
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: viktoriya.kozyreva@gmail.com

Abstract[1]

In this paper we consider problem of analysis and comparison of curricula and academic programs for Computer Science majors of different foreign and domestic universities. The main difficulties there are that curricula are presented in documents of different structure and format. Each country and sometimes university has own way to form curricular. We investigate the possibility of constructing an information system based on the expert system technology for such analysis and comparison. Additionally, in this paper we analyse the sequence of steps that must be taken for solving the comparison problem.

1. Introduction

Comparative analysis of syllabi of different countries can detect useful features and characteristics of educational process, which can be applied while modernization of Russian educational system. Besides that, academic internalization is one of educational development directions. Base for measuring and comparison of the education programs is comparative analysis. In this work we concentrate on the undergraduate academic programs in Computer Science. Comparative analysis of the curricula and academic programs is a labour intensive procedure; therefore there exists a need for creation of an automatic information system for this task. There are three types of tasks for which such information systems are used for: structured (formalizable) tasks, unstructured (unformalizable) tasks, and partially structured tasks.

In real life application there are only few fully structured and fully unstructured task and usually for a given task only some elements and relations between them are known. These tasks are called partially structured. Systems that handle these tasks would use help of a human in the analysis of the information. However, these systems are called automated systems, since humans participate in their operation. Comparison of the academic curricula is an example of such partially structured task, since the academic programs can be compared partially they have at least some matching of the names of the courses listed. In the worst case, when the names of the courses don't match completely the analysis is extremely hard. Information technologies used in the expert systems provide high-level support of the decision making process. In this work we will consider the models of knowledge representation and algorithms for comparison of academic programs.

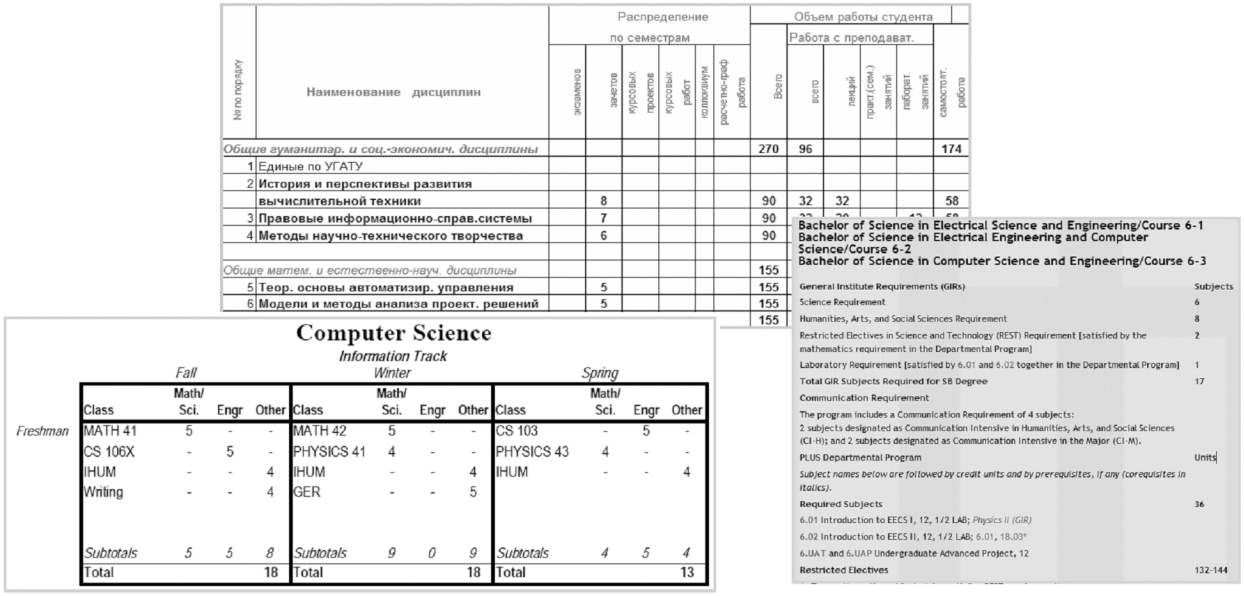
2. Problem of curricula comparative analysis

While comparison of curriculas following difficulties appear:

1. Curriculas of different countries are presented in different documents format (PDF, HTML). There is no one structure for curricula. Examples of curricula's fragments presented on Figure 1.
2. Different units of academic load: for example, in the U.S. academic load is measured in units, in Russia – in academic hours: 1 unit is 3 hours of work per week for one semester. Semester consists of 10 weeks, so 1 unit equals to 30 academic hours.
3. Different systems of knowledge evaluation.

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

4. Educational programs are regulated by different structures and organizations. Educational program



№ по порядку	Наименование дисциплин	Распределение по семестрам: экзаменов	зачетов	курсовых проектов	курсовых работ	контрольных	расчетно-граф работа	Объем работы студента: Всего	Работа с преподават.: всего	лекций	практ (сем.) занятий	лаборат. занятий	самостоят. работа
Общие гуманитар. и соц.-экономич. дисциплины								270	96				174
1	Единые по УГАТУ												
2	История и перспективы развития вычислительной техники		8					90	32	32			58
3	Правовые информационно-справ.системы		7					90					
4	Методы научно-технического творчества		6					90					
Общие матем. и естественно-науч. дисциплины								155					
5	Теор. основы автоматизир. управления		5					155					
6	Модели и методы анализа проект. решений		5					155					
								155					

Computer Science

Information Track

	Fall				Winter				Spring			
	Class	Math/ Sci.	Engr	Other	Class	Math/ Sci.	Engr	Other	Class	Math/ Sci.	Engr	Other
Freshman	MATH 41	5	-	-	MATH 42	5	-	-	CS 103	-	5	-
	CS 106X	-	5	-	PHYSICS 41	4	-	-	PHYSICS 43	4	-	-
	IHUM	-	-	4	IHUM	-	-	4	IHUM	-	-	4
	Writing	-	-	4	GER	-	-	5				
	Subtotals	*5*	*5*	*8*	*Subtotals*	*9*	*0*	*9*	*Subtotals*	*4*	*5*	*4*
	Total			18	Total			18	Total			13

Bachelor of Science in Electrical Science and Engineering/Course 6-1
Bachelor of Science in Electrical Engineering and Computer Science/Course 6-2
Bachelor of Science in Computer Science and Engineering/Course 6-3

General Institute Requirements (GIRs)	Subjects
Science Requirement	6
Humanities, Arts, and Social Sciences Requirement	8
Restricted Electives in Science and Technology (REST) Requirement [satisfied by the mathematics requirement in the Departmental Program]	2
Laboratory Requirement [satisfied by 6.01 and 6.02 together in the Departmental Program]	1
Total GIR Subjects Required for SB Degree	17
Communication Requirement	
The program includes a Communication Requirement of 4 subjects: 2 subjects designated as Communication Intensive in Humanities, Arts, and Social Sciences (CI-H); and 2 subjects designated as Communication Intensive in the Major (CI-M).	
PLUS Departmental Program	Units
Subject names below are followed by credit units and by prerequisites, if any (corequisites in italics).	
Required Subjects	36
6.01 Introduction to EECS I, 12, 1/2 LAB; *Physics II (GIR)*	
6.02 Introduction to EECS II, 12, 1/2 LAB; 6.01, 18.03*	
6.UAT and 6.UAP Undergraduate Advanced Project, 12	
Restricted Electives	132-144

Figure 1. Fragments of Stanford (USA), USATU (Russia), MIT (USA) syllabi

Stanford University		Ufa State Aviation Technical University	
Discipline	Hours	Hours	Discipline
Calculus	300	340	Mathematical analysis
Integral Calculus of Several Variables	150		
Mathematical Foundations of Computing	150	140	Discrete Mathematics
Linear algebra and matrix theory	150	140	Algebra and geometry
Applied Number Theory and Field Theory	150		

Figure 2. Comparison of disciplines (block "Mathematics")

can vary within one country. For example educational programs in the U.S. are regulated by university and an accreditation program and may differ from each other more than Russian programs.

5. Several subjects of one plan can meet one subject of another plan. For example, subjects «Calculus» and «Integral Calculus of Several Variables» of Stanford University curriculum relevant to discipline "Mathematical analysis" of USATU curriculum (Figure 1).
6. Different loads for the same subjects. For example, the number of hours, allocated to the subject "Discrete Mathematics " in USATU, equals 140. In Stanford University to study the same subject, «Mathematical Foundations of Computing», you need 150 hours (Figure 1).

The analysis of the curricula will allow us to identify relations, patterns and similarities of the programs and store this knowledge in the knowledgebase.

The key step in any work that deals with knowledge is the forming of the knowledge field. This is a non-trivial task that involves exposure and identification of the objects and concepts in the data domain, their properties and relationships between them, as well as their representation in a direct and obvious form. Without careful structurization of the field of knowledge the creation of the knowledgebase is impossible.

Databases are usually used for storing the information. The databases are characterized by the large volume of the information that they can store and relatively cheap average cost of storing the information. The knowledge is usually stored in knowledgebases which have small volume, but extremely expensive information collection.The knowledgebase is the foundation of any intelligent system. The knowledge in it is described on a particular language that is very similar to one of the natural languages.There exist several models for representation of problems for different data domains. Most of them, however, belong to the four of the following classes: Production models; Semantic networks; Frames; Formal logical models [1].

There are several information systems being developed for conducting the comparative analysis of the academic programs. All of them are based on the expert system technology. Some of them are used a combination of network and production models. In these systems the declarative knowledge is described in the network component, while the procedural knowledge is described in production element. In this case the approach is referred to as production model operating on a semantic network.

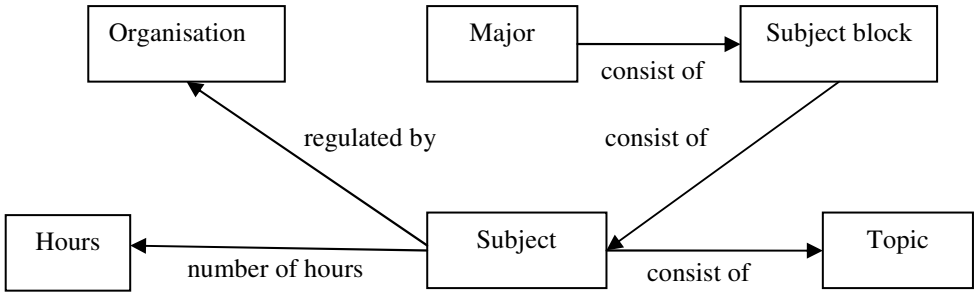


Figure 3. Ontology fragment for curriculum

3. Algorithms for comparison of the academic curricula

In the scope of this project we also conducted analysis of data domain and constructed an algorithm for the academic program comparison process:
Step 1: Identify the subjects that are similar by topics covered. First of all during this step the initial data required for program comparison is inputted into the system by the user. Then the facts and thesaurus are examined in attempt to identify subjects in the compared plans that might be similar based on the title or keywords in the description. The comparison done in this step is fully automatic and it is conducted by using semantic network.
Step 2: Comparison of subjects that were not matched with others in automatic mode. During the first step the systems keeps a list of subjects that for some reason were not matched with courses on the other academic program. This list is then given to the human expert, who has the following options for matching these courses:
Matching by title. In this case the foreign course is automatically grouped with some subject block.
Matching by keywords. In this case the foreign course is matched to the Russian academic program subjects based on the description keywords. The foreign course is also linked by title.
Matching by title to a subject group. This action makes sense only when expert doesn't know which courses can match this subject exactly, but knows which subject block the course in question belongs to.
If an expert cannot match the subject to a block or another subject this course is moved out to the block of unmatched subjects. In this case the system would report an assumption that there exist subject that cannot be compared with others.
Step3: Compare the academic course load of the matched subjects. During this phase the system analyses the lists of the comparable courses and calculates the course load for them. The comparison is done based on the assumption that the course load is divided equally by the number of matching if there are more than one courses matched with this one.

The algorithms for comparison of the academic curricula reflect the sequence of actions taken by an academic process expert, who conducts a heuristic and linguistic analysis of the contents of two academic programs.

4. Knowledge representation models

A semantic network is a set of related expressions (words and word combinations). The semantic networks usually contain words that are most often used in the text and have substantial semantic meaning. For each expression the networks forms a set of associative relations (relations by meaning), i.e. the list of other expressions, which were used in combination with this expression in the text [1].
During the analysis of the text one can employ the semantic network built from other (reference) texts. For example, given a semantic network for academic program for major 230100 "Informatics and computer science", we can use it for filtration of the information in other academic programs from different foreign universities. In this case only the subjects contained in the reference network will be identified in the academic curricula under investigation, and the resumes of the curricula would be built based on these subject. The comparison of the semantic networks of different academic programs allows establishing the degree of their similarity, which can be used for automatic structurization.
Describe a fragment of the semantic network presented on Figure 3 by example the academic curriculum of the major 230100 "Informatics and computer science" for math and science subject blocks (MSS). The network describes the structure of the curriculum in terms of concepts and relations between them. We can also represent the above network in words as follows: A particular *major (230100) consists of*_several *subject blocks*, *for example MSS* block. Each of the *subject blocks consists of* several *subjects*. *For example*, subjects of *mathematical analysis, linear algebra, geometry, automation theory, physics, ecology, discrete math, probability theory and statistics* belong to the *MSS* block. Each *subject* is characterized

by its attributes *Number of hours* and *organisation/component,* which it is regulated by (for example, *regional component*), and is described by a set of *keywords/topics* (for example, for subject *"Discrete Math"keywords are graph theory, set theory).*

4. Conclusion

In this paper we considered the problem of curricula analysis and comparison. We investigated using a semantic network for structuring and formalization of the knowledge and their representation in an expert system for comparative analysis of academic programs of different foreign and domestic universities. We gave an example of the algorithm for subject matching and comparison.

References

1. A.P. Chastikov, T.A., Gavrilova, D.L. Belov. Development of expert systems. CLIPS environment. – St.P.: BHV-Petersburg, 2003.
2. National academic standard for major 552800 (230100) "Informatics and Computer Science".
3. Stanford University. URL: www.stanford.edu (last access: 01.02.2010).

A branch-and-price method for the 1D contiguous bin packing problem and its application

Marat Mesyagutov
Dresden University of Technology,
Institute for Numerical Mathematics
Germany
e-mail: mmesyagutov@gmail.com

Abstract[1]

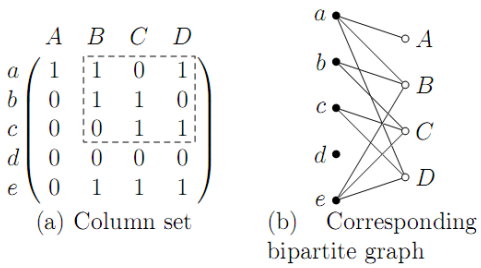
The 1D contiguous bin packing problem is considered. A branch-and-price algorithm for its solution based on combinatorial characterization of matrices having contiguous property is proposed.
Key words: 1D contiguous bin packing problem, branch-and-price

1. Introduction

Let us consider the one-dimensional contiguous bin packing problem (1CBPP). Given m item types with lengths $w=(w_1,\dots,w_m)$ and quantity $b=(b_1,\dots,b_m)$ for each type $i\in I=\{1,\dots,m\}$, the items should be packed into one-dimensional containers (bins) with length W. The order of bins is fixed. The packing has to satisfy the following restrictions:

1. The total length of items packed into a bin must not exceed the length of the bin.
2. The total amount of an item packed into a bin is less or equal to 1.
3. Each item type is packed contiguously (C-property), i.e. if an item is packed into bins j_0 and j_1 ($j_0 < j_1$) then it is also packed in all bins j with $j_0 < j < j_1$.

Fig. 1 (a) shows the column set (set of packings) which does not satisfy the third restriction of contiguous packing feasibility, despite that set $\{A,B,C\}$ of columns has the C-property.
The 1CBPP has a lot of applications, e.g., in scheduling [1]. The problem arises as a relaxation of the 2D rectangular strip packing problem (2SPP) [2].



(a) Column set (b) Corresponding bipartite graph

Figure 1. The column set does not have the C-property. The property is broken in the subcolumns of columns B, C, D, which are marked by the dashed box.

2. Formal problem formulation

Let $A=\{a^j\in\{0,1\}^m : w^T a^j \le W,\ j\in J\}$ with $J=\{1,\dots,n\}$ be the set of all possible ways of bin packing, where $a_i^j=1$, if item i is packed into a bin according to packing manner j. Formally, the problem can be described as follows:

$$\sum_{j\in J} x_j \to \min; \quad s.t. \tag{1}$$

$$\sum_{j\in J} a^j x_j = b; \tag{2}$$

$$x_j \in \mathrm{Z}_+,\ \forall j\in J; \tag{3}$$

$$a^j \text{ with } x_j > 0 \text{ form a contiguous solution} \tag{4}$$

Constraint (4) means, that the subset of columns has to have C-property (or can be permuted, so that the resulting set will have the C-property).
In order to satisfy constraint (4) branch-and-price algorithms are considered, which use sets of columns to define subproblems.

3. Characterization of column sets having the C-property and algorithms

Definition 1. Let $C=\{c^j\in\{0,1\}^m : j\in J_C\}$ be a column set. The bipartite graph $B_C=(V_1,V_2,E_C)$

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

associated to column set C is defined by $V_1 = I$, $V_2 = J_C$, and $E_C = \{(i, j) : i \in V_1, j \in V_2, c_i^j = 1\}$.

In Fig. 1 (b) is shown the bipartite graph corresponding to the set of columns in Fig. 1 (a).
The characterization of graphs corresponding to column sets having C-property was given by Tucker [3], and it is based on a characterization in terms of so-called *asteroidal triples*.

Definition 2. Let $G = (V, E)$ be a graph. Three vertexes from V form an asteroidal triple if between two of them there exist a path in G that does not contain any vertex from the closed neighborhood of the third vertex.
The following statement is true.

Theorem 1 ([3]). Let $B_C = (V_1, V_2, E_C)$ be the bipartite graph associated to column set C. Set C has the C-property if and only if V_2 contains no asteroidal triple of B_C.

A set of columns is said to be C-permutable if there exists a permutation of its columns, which has the C-property. It is also true the following corollary of the Theorem 1.

Theorem 2. Let C^+ be a column set having the C-property and $c^k \in \{0,1\}^m$ be a column with $c^k \notin C^+$. Column set $\tilde{C} = C^+ \cup \{c^k\}$ is not C-permutable if and only if bipartite graph $B_{\tilde{C}} = (I, \{1, \ldots, |\tilde{C}|\}, E_{\tilde{C}})$ has at least one asteroidal triple.

Using Theorem 2 we propose exact algorithms based on branch-and-bound method. Two main procedures which should be defined are branching and bounding. Branchings consist in consecutive building a feasible 1D partial packing having contiguous property. On each step we have a set of fixed columns, which are in the packing and a set of forbidden columns, which must not appear in the packing. Herewith, a branching tree is build, where each node is characterized by these two sets.

Lower bound (bounding procedure) for a node is calculated by solving optimally the continuous relaxation of problem (1)-(3) with an altered right hand side $b - \sum_{c^j \in C^+} c^j$, i.e. with relaxed condition (3): $x_j \in \Re_+$. Since A is an implicitly defined column set we apply the column generation method, where forbidden columns are not allowed to be generated.

In order to tighten the lower bound and generate less descendants we forbid not only complete columns but also sub-columns, which break the C-property of the set of fixed columns.
Let subcolumn d^*, which breaks the C-property of C^+, be given by index set $I_* \subset I$ and coefficients $d_i *$, $i \in I_*$. Let $C(I_r, I_c) = \{(c_i^j) : i \in I_r, j \in I_c\}$ be the set of columns from $I_c \subset J_C$ and rows from $I_r \subset I$. If $C^+(I_*, J_C) \cup \{d^*\}$ does not have the C-property, then d^* is called ineligible. An ineligible subcolumn is called minimal if does not exist $i^* \in I_*$: $\bar{d}^* = d_i$ if $i \in I_* \setminus \{i^*\}$, and $\bar{d}^* = 1 - d_i$ if $i = i^*$, so that $C^+(I_*, J_C) \cup \{\bar{d}^*\}$ does not have the C-property.

Theorem 3. Let d^* be a minimal ineligible subcolumn. Then $1 \le \sum_{i \in I_*} d_i * \le 3$ is true.

Theorem 4. Let d^* be a minimal ineligible subcolumn. If there is an asteroidal triple containing the vertex corresponding to d^*, then $1 \le \sum_{i \in I_*} d_i * \le 2$ is true.
Using characterization given in Theorems 3 and 4 we develop algorithms which find ineligible subcolumns. There exist two possible strategies. The first one is to find all ineligible subcolumns of a column which breaks the C-property. The second strategy is to find all ineligible subcolumns of a given set of columns having the C-property. Both strategies differ by complexity of realization algorithms.
Results of numerical experiments will be presented on test instances from the literature.

Acknowledgements

The project was supported by the Erasmus Mundus ECW.

References

1. S. Hartmann. Project Scheduling under limited resources. *Models, methods and applications*. Springer, Berlin, 1999.
2. S. Martello, M. Monachi, D. Vigo. An exact approach to the strip-packing problem. *INFORMS Journal on Computing*. 15(3):310--319, 2003.
3. A.C. Tucker. A structure theorem for the consecutive 1's property. *Journal of Combinatorial Theory, Series B*. 12:153--162, 1972.

Novel Technologies in Industrial Tasks on example of Bin-Picking

Ilshat Mamaev
Computer Science and Robotics Department
Ufa State Aviation Technical University
Ufa, Russia
e-mail: ilshat.mamaev@gmail.com

Igor Tchouchenkov
Institute for Process Control and Robotics (IPR)
University of Karlsruhe
Karlsruhe, Germany
e-mail: igor.tchouchenkov@kit.edu

Nafisa Yusupova
Computer Science and Robotics Department
Ufa State Aviation Technical University
Ufa, Russia
e-mail: yussupova@ugatu.ac.ru

Heinz Wörn
Institute for Process Control and Robotics (IPR)
University of Karlsruhe
Karlsruhe, Germany
e-mail: woern@kit.edu

Abstract[1]

Emerging technologies allow proposing new solutions for classical problems in Robotics. This article addresses a bin-picking problem. In this task robot is required to grip an object from a stack of scattered objects placed in a bin. To perform such tasks, a vision system is essential. In recent time, a lot of solutions based on 2D- and 3D-vision have been developed. But in spite of this there exist several problems that can be solved more efficiently using a novel PMD (Photonic Mixer Device) sensor technology, e.g. robust and fast scene reconstruction and analysis. The main feature of this technology that can help to increase effectiveness of the bin-picking system is the possibility to obtain simultaneously depth and amplitude images in real-time.
However, PMD has some drawbacks which are discussed and approaches to overcome them are proposed. As result, an extensible modular framework for solving bin-picking problem in real-time is presented. The developed system was implemented and validated for bin-picking task on a KUKA KR 60 HA robot.

1. Introduction

One of the main tasks for industrial robots is assembly operations and pick and place operations, where robots can handle only objects with known locations and orientations. The robotic system without sensors is used only for preprogrammed repetitive tasks.

However, over the last years, the bin-picking problem has not received as much attention, in spite of the absence of general solution. At that time, due to sufficiently large enough line volumes in manufacturing industry, singulation, fixturing and binning were more efficient and effective that moving forward with a vision based bin-picking solution. Today, however, the manufacturing industry is facing a major paradigm shift. For example in automotive industry introducing a new model of automobile generally takes three to five years from inception to assembly. Ideas for new models are developed to respond to unmet pubic needs and preferences. Redesigned and restyled models appear on the streets almost each one or two years. Production cycles are getting shorter and shorter, approaching number of months rather than years. This change puts a great deal of pressure on any kind of fixed manufacturing lines – flexible manufacturing is again come to the fore for industry.
Meanwhile, new sensor technologies, advances in processing speed, new algorithms, and significant engineering developments have been developed, which allow to propose new solutions for the bin-picking problem.
In this paper we present a new approach to bin-picking problem solution that is based on PMD-sensors and propose real-time framework that can help us to achieve promising results.
In section 2 bin-picking problem is described and existing solutions based on different principles are discussed. In section 3 a novel PMD-technology and its advantages and disadvantages are presented. The general description of real-time framework for bin-picking problem is given in section 4. Conclusion and

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

future work are presented in the last section of this paper.

2. Bin-Picking Problem Overview

As it was mentioned before the bin-picking problem is one of the well-known problems in robot vision. It contains the following tasks:

- 3D scene reconstruction;
- scene analysis that consists of:
  - object recognition: recognition of the ungraded objects of the different classes that can be occluded and overlapped and have undefined position;
  - object localization: recognition of the objects positions in space and it's orientation;
- calculation of the gripping points of the recognized objects;
- developing the extraction strategy for the recognized objects;
- path planning: path planning for grasping the recognized objects taking into consideration the collision avoidance.

Bin-picking system consists of the software and hardware modules. It is necessary to mention that among all hardware modules sensor system is of large importance. On the selection of sensor system depend the speed and accuracy of 3D reconstruction that has an influence on scene analysis and path planning processes.

All sensor systems can be classified in 2D- and 3D-solutions. 2D-systems are based on monocular cameras (monochrome or color). 3D-systems are based on laser range finders, stereo cameras, triangulation systems etc. [1] All these systems have a number of disadvantages, such as [2]:

- strong dependence on lighting conditions, on the colour and on the reflectance features of the objects (monocular and stereo cameras);
- low speed of the image acquisition (laser range finders);
- high complexity and low speed of the 3D scene reconstruction (stereo cameras);
- expensive solutions;
- large-sized devices.

In recent years a novel PMD-technology is being developed that has some features which allows to achieve better effectiveness in solving bin-picking problem. This technology and its features are described in next section.

3. PMD-Technology

3D sensor systems have been investigated for several decades. In recent years, improvements on classical approaches like stereo vision and novel techniques have emerged, leading to 3D vision systems with radically improved characteristics.

Figure 1. Ifm o3d PMD-camera.

A rather new and promising approach developed during the last years is to estimate distance by time-of-flight measurements based on Photonic Mixer Device (PMD) Technology. To observe a scene so-called PMD-cameras emit modulated infrared light using LEDs with 20 MHz modulation frequency and measure the time between emitting the modulated light and receiving the echo of the light using a special PMD correlation sensor element. In PMD sensor each pixel can individually measure phase shift of the modulated light. Detailed description of physics basic principles can be found in [3, 4].

This new technique delivers depth images at a resolution up to 120x160 pixels at frame rates up to 25Hz (ifm o3d PMD-camera which is presented in Figure 1)

Figure 2. Real objects (left), intensity (middle) and depth (right) images sample.

The cameras are suitable for near range measurements of up to 7.5 m distance and potentially offer dense depth at no additional computational costs [5]. The cameras offer both a reflectance image with IR modulation intensities (so-called intensity image, see Figure 2) and a range image with pixel-wise depth (so-called depth image, see Figure 2).

Unlike other systems, the PMD-system is a very compact device which fulfills the above stated features desired for real-time distance acquisition and, due to an automatic suppression of background light, suitable for indoor as well as outdoor dynamic scenes.

Nevertheless, there are shortcomings which have been solved in the framework:

- *low resolution and small field-of-view (FOV)*: these drawbacks complicate such important tasks in bin-picking problem as recognition and 3D scene reconstruction. To avoid this we

propose to use scene reconstructed from multiple images (real-time image acquisition allows it);

- *distortions and depth measurement error*: to take into account the precision requirements of bin-picking we have manage this with the aid of special calibration methods;
- *noisy images*: noise reduction filters (mean, smoothing, oriented median [3]) are used for elimination that.

4. Real-time Framework

4.1. Scene reconstruction

The possibility to obtain simultaneously 2,5D depth and amplitude images in real-time and the small size of the camera are two important advantages among others.

The acquisition of 2,5D images in real-time gives us possibility to carry out robust and fast scene reconstruction. There is no need to use complicated and computationally intensive methods for it. 3D scene reconstruction algorithm can be realized as following (see also Figure 3) [2]:

1. Depth images are being transformed in 2,5D models into camera coordinate system. Each pixel of depth image (p_{img}) is being transformed into space point in camera coordinate system (p_{camera}) using projection matrix (M_{proj}):

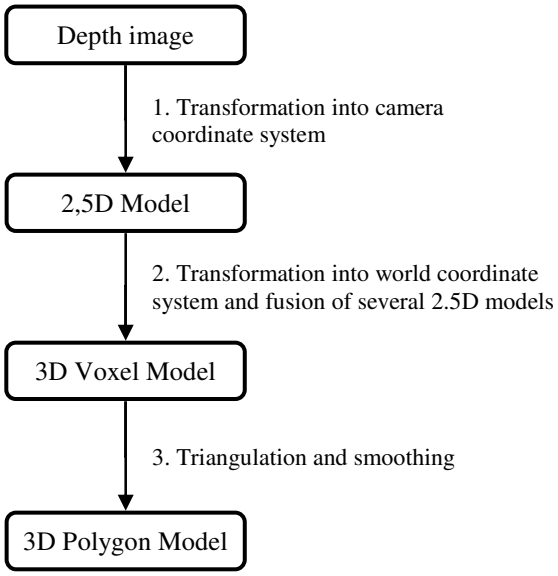


Figure 3. Scene reconstruction algorithm

$$p_{img} = M_{proj} \cdot p_{camera}, \quad (1)$$

$$\lambda \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad (2)$$

$$\begin{cases} u = \dfrac{f_x \cdot x + s \cdot y + c_x \cdot z}{z} \\ v = \dfrac{f_y \cdot y + c_y \cdot z}{z} \end{cases} \Rightarrow \begin{cases} x = \dfrac{(u - c_x) \cdot z}{f_x} - y \\ y = \dfrac{(u - c_y) \cdot z}{f_y} \end{cases}, \quad (3)$$

where λ – scale factor; f_x, f_y - focal lengths expressed in pixel-related units; c_x, c_y - principal point (image center); u, v - coordinates of point in pixels; x, y, z - coordinates of a 3D point in space.

2. 2,5D models are being transformed into world coordinate system:

$$p_{world} = {}^{world}T_{TCP} \cdot {}^{TCP}T_{camera} \cdot p_{camera}, \quad (4)$$

where ${}^{TCP}T_{camera}$ - transformation matrix from camera coordinate system into coordinate system of TCP (this matrix can be calculated after carrying out hand-eye-calibration). ${}^{world}T_{TCP}$ - transformation matrix from coordinate system of TCP into world coordinate system (this matrix can be calculated using known robot position).

3. 3D voxel model is being transformed into polygon model using Delaunay triangulation and then smoothed using Laplace filter.

The sample of the 3D reconstruction is presented on Figure 4.

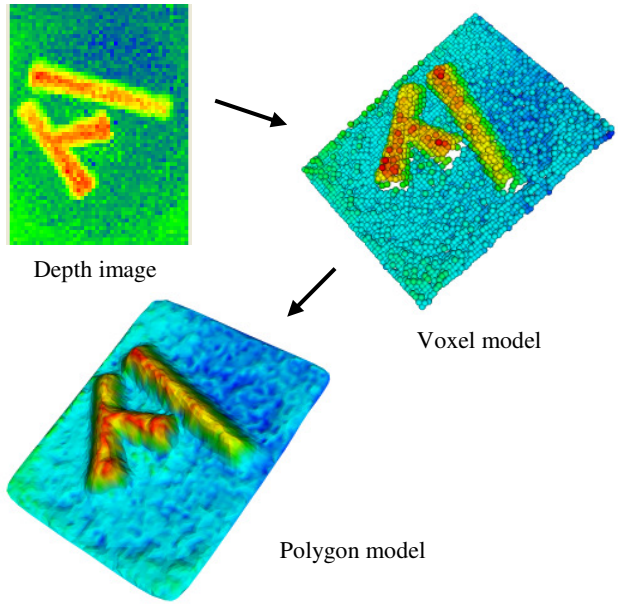


Figure 4. Scene reconstruction sample

4.2. Framework description

4.2.1. Framework structure

Class diagram of the proposed real-time framework for bin-picking problem is presented on Figure 11. This framework consists of the following modules:

1. ***Main control module***: fulfils network communication management, processes XML messages from modules taking into account rules of resources access policy, stores framework service information, etc.;
2. ***Robot control module***: carries out robots control (several robots simultaneously) using XIRP protocol;
3. ***PMD-camera control module***: synchronizes images from several PMD-cameras and stores them, controls camera parameters;
4. ***Environment modeling module***: reconstructs scene using special algorithms, calculates voxel and polygon scene models;
5. ***Object recognition module***: recognizes objects and their location in space using hierarchical iterative object recognition method [6, 7, 2];
6. ***Camera calibration module***: provides automated lateral, hand-eye and depth calibration of PMD-cameras;
7. ***Path planning module***: calculates path without collisions for each robot movement;
8. ***Extraction strategy module***: works out the strategy of objects extraction (order, priority, etc.);
9. ***Gripping points calculation module***: calculates gripping points for extracting object.

4.2.2. Communication

Communication between modules is carried out using XML-based protocol. An XML-scheme of a message is presented on Figure 5.

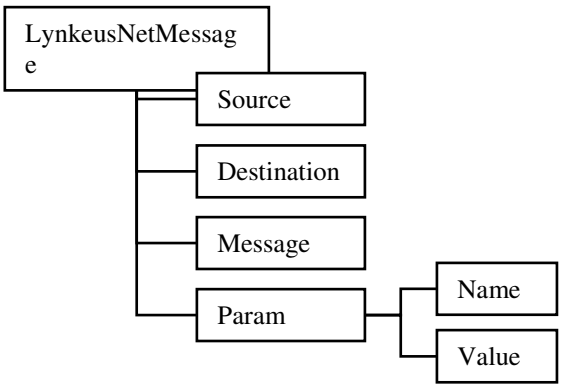


Figure 5. XML-scheme of the network message

An XML message can be presented in the following way:

```
<LynkeusNetMessage>
    <Source>ENVIRONMENT_MODELING_MODULE
    </Source>
    <Destination>PMD_CAMERA_CONTROL_MODULE
    </Destination>
        <Message>GET_IMAGE</Message>
        <Param>
                <Name>ImageType</Name>
                <Value>Depth</Value>
        </Param>
</LynkeusNetMessage>
```

Elements of this message:

<LynkeusNetMessage> - header-tag;

<Source> - contains the name of the module which sent the message;

<Destination> - contains the name of the receiver module;

<Message> - defines the message (GET_IMAGE as example, message list presented below);

<Param> - message parameters. Multi-parameter format is accepted. Each parameter consists of name and value;

<Name> - parameter name;

<Value> - parameter value.

4.2.3. Module architecture

Each module of the framework is realized as a single process or thread that provides XML-interface for intercommunication with other module or with GUI-application (see Figure 6).

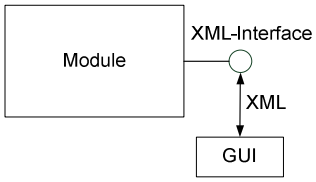


Figure 6. Module-GUI interface

In one's turn each module can consist of sub-modules as threads (see Figure 7). One sub-module acts as a control module organizing work of other sub-modules and communicating with other framework modules.

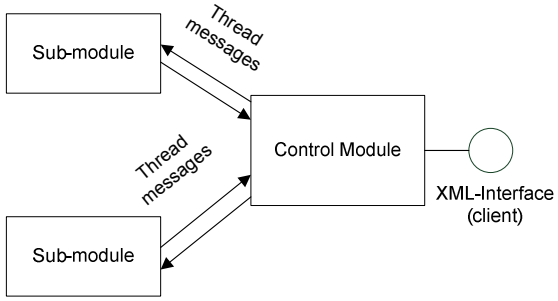


Figure 7. Module-submodules interface

4.2.4. Framework architecture

It is possible to develop the framework using two main architecture principles:

- Client-server architecture (see Figure 8);
- Multi-agent architecture (see Figure 9).

The main difference between these two architectures is the way of modules intercommunication. In first case (client-server) all communications are being held through main control module that grants or denies the access to the resources.

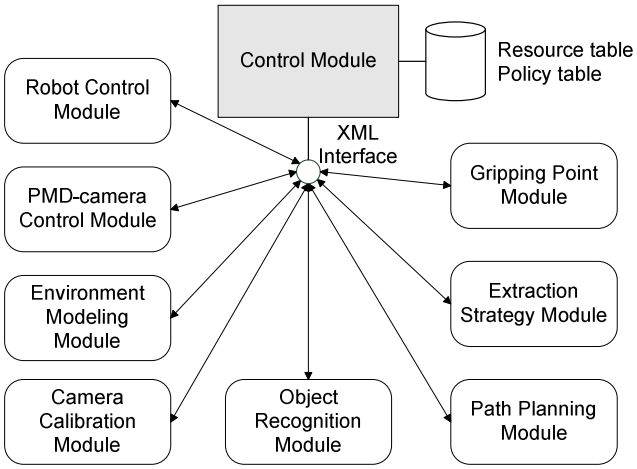


Figure 8. Client-server architecture

In the second case (multi-agent architecture) each module can communicate with other modules directly without using main control module. In this case each module must control its resources by itself.

Framework was developed in a way that allows to organize intermodule communication using both architectures. Effectiveness of such system depends on specific task.

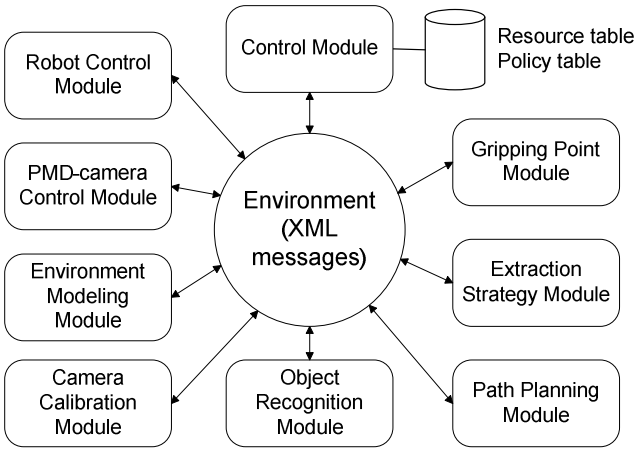


Figure 9. Multi-agent architecture

5. Conclusion and future work

In this paper we have proposed and presented real-time framework for solving bin-picking problem using novel PMD-technology. On Figure 10 the real and virtual scenes of the developed bin-picking system is shown.

In the nearest future we would like to continue our development of the bin-picking system, to carry out more experiments.

Another research field is development and implementation of more sophisticated algorithms for optimizing collision-free path planning.

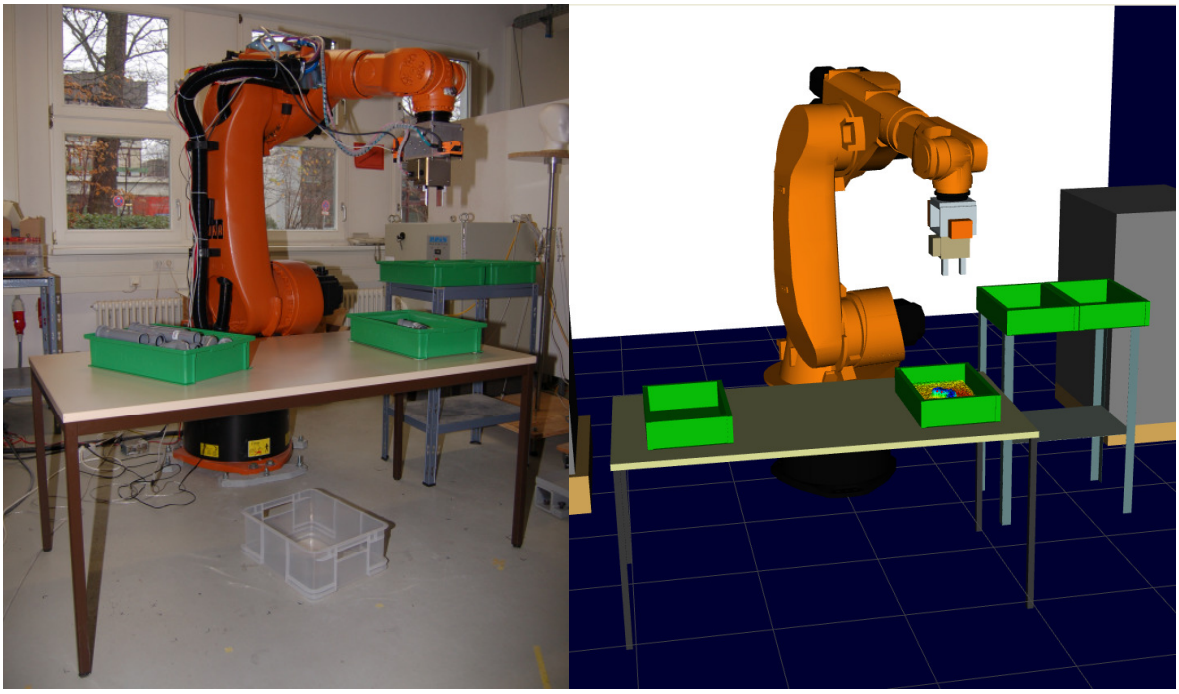


Figure 10. Real and virtual scenes of the bin-picking system.

Acknowledgement

This work is partially funded by grant 16SV2307 from the German Federal Ministry of Education and Research (BMB+F).

References

1. M. Müller, “Picking strategies for object singularization using industrial robots,” Ph.D. thesis, Universität Karlsruhe (TH), 2001.
2. K. Safronov, I. Tchouchenkov, and H. Wörn, “Hierarchical Iterative Pattern Recognition Method for Solving Bin Picking Problem”, in Proc. Robotik 2008.
3. J. Graf and H. Woern, ”Oriented Median Filtering Preserving Significant Image Structures”, in Proc. 9th Int. Workshop on Computer Science and Information Technologies (CSIT), vol. 2, Ufa, Russia, 2007, pp. 4-8.
4. B. Buxbaum, R. Schwarte, T. Ringbeck, M. Grothof, X. Luan, "MSM-PMD as correlation receiver in a new 3D-ranging system", SPIE – Remote Sensing, Laser Radar Techniques: "Ranging and Atmospheric Lidar", Toulouse 2001.
5. H. Kraft, J. Frey, T. Moeller, M. Albrecht, M. Grothof, B. Schink, H. Hess and B. Buxbaum, 2004. 3d-camera of high 3d-framerate, depth-resolution and background light elimination based on improved pmd (photonic mixer device)-technologies. In: 6th International Conference for Optical Technologies, Optical Sensors and Measuring Techniques (OPTO 2004).
6. J. Graf, K. Safronov, I. Tchouchenkov, and H. Wörn, “Automated registration of medical head volume images using gradient based method,” in Proc. 7th Int. Workshop on Computer Science and Information Technologies (CSIT), vol. 1, 2005, pp. 174–179.
7. K. Safronov, I. Tchouchenkov, and H. Wörn, “Optical character recognition using optimisation algorithms,” in Proc. 9th Int. Workshop on Computer Science and Information Technologies (CSIT), vol. 1, 2007, pp. 85–89.

Identification Methods Usage at Various Stages of Technical Systems Lifecycle

I. A. Krivosheev
Ufa state aviation technical University I
Ufa, Russia
e-mail: krivosh@sci.ugatu.ac.ru

Abstract[1]

Basic aspects of composing of simulation models of technical systems are examined. Identification methods are considered for various stages of technical systems design, finishing, and operation, for diagnostics of their technical condition. Special features of simulation models identification are shown. Examples of identification and usage of models of gas turbine engines and power plants are given. Special features of obtaining and using of gas turbine engine operation maps are examined, for gas turbine engine units and parts as well. Necessity and possibility of presenting operation maps in a maximally broad parameter range is shown. New methods for operation maps representation are proposed, in generalized dimensionless form.The text of your abstract should be placed here. It is in your own interest to ensure the abstract adequately describes the content of your paper.

1. Introduction

At extent of a technical system lifecycle, beginning from design and ending with operation, an evolving mathematical model accompanies it – in accordance with computer aided lifecycle support concept (CALS). For maintaining the integrity of united information space, in the process of a technical system finishing and operation, it is necessary for models of different level to interact between each other, according to the results of tests and monitoring. The effectiveness of design, finishing, and fixing of gas turbine engine, monitoring of its technical state in the operation, is determined by degree of adequacy and productivity of models utilized in these stages. At this statistical models of general type of engines are distinguished from individual mathematical models for given engine instance. The main problem of model identification in the stage of finishing and operation, and state monitoring of a technical system, is experimental data lack. In connection with this, authors are proposed special methods and means, based on imitation net simulation, as in [8].

The family of models can be represented by a set of program modules, connected to each other via common database. The model of a make, represented in database, can have hierarchical or net structure. Usually the version of information lifecycle support is used, where each model corresponds to a given design-and-finishing procedure (within the framework of corresponding lifecycle stage). This is determined, first of all, by nomenclature of input (X) and output (Y) model data. This type of organization is called procedural. Program modules simulate design-and-finishing procedures in it.

Contemporary objective approach requires that the program modules would simulate object itself – a technical system and its structural elements - in specific aspect, taking into account specific factors. This approach is implemented in contemporary systems of imitation simulation, such as developed with participation of the author CAMCTO system and CAMCTO-made application DVIG [2,4], Fig.1.

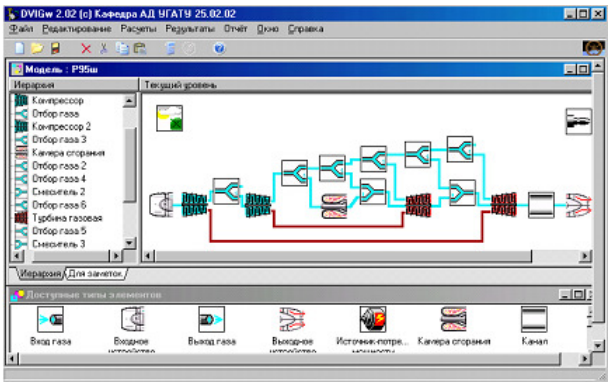


Figure 1. DVIG Modelling System.

In this case, the mathematical models of structural elements are formed and developed in the process of the design of technical system, in accordance with the methodology proposed, and using the developed means. In particular, during the first stage of gas turbine engine design, structural elements are used whose models reflect thermo-gas-dynamic aspect. Considering this, "ports" and information "flows" are

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

provided of type "GAS", "MECHANICS", etc. Internal algorithms and nomenclature of the parameters of structural elements are based on theory of air engines, thermal design of boilers, theory of steam turbines, calculations of the fundamental thermal circuit of steam-turbine installations, and theory of heat pumps.

2. Obtaining operation maps of air engine units, based on model identification

The method is developed [7,8], which makes it possible to improve the process of the identification of the simulation model of engine. The method is able to solve problems with experimental data lack - by sequential changing during identification the model structure, and acquisition of additional information, taking into account the degree of its authenticity. The method includes the series of the stages:

- Bringing the **model structure** into correspondence with the structure of experimental information and step by step model change in the process of the identification;
- Analysis of **correctness of the tasks**, solved in each stage of identification with aid of the network, which depicts an interrelation of the parameters;
- Formation of corresponding **intermediate models**;
- Giving **conditions** of conducting the **identification** (composing calculation laws);
- Giving experimental values of parameters in the model (in the form of functions and tables);
- Obtaining (by iteration technique, on the basis of the principle of greatest likelihood) functions of corrections for **dimensioning parameters** for unit **operation maps**;
- Introduction of functions of corrections into the model, refinement of the existing operation maps;
- **Obtaining missing operation maps** using the proposed method;
- Consecutive including of obtained dependences into the investigated model during identification.

3. Unification of operation maps of gas turbine engine units

Operation map is the mean of agreement between models of different level. Therefore, for each structural element it is necessary to obtain necessary and sufficient collection of characteristics. It is shown that operation maps represented in normalized ("dimensionless") form provides the possibility of agreement of structural modeling elements in the composition of simulation model, including cases of identification using experimental data.

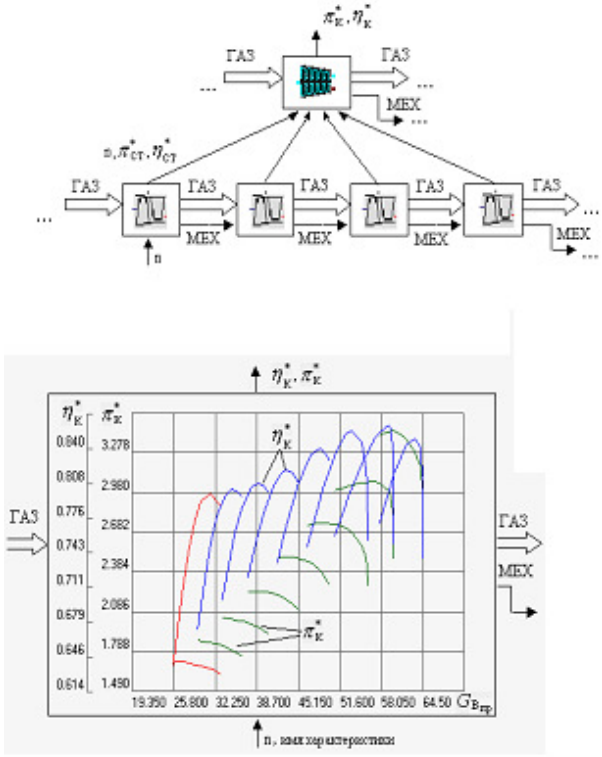


Figure 2. Information Streams Between Elements and Parameters in an Engine Model, their Net Representation Method (Based on Example of Compressor)

During solution of such identification tasks of gas turbine engine models, the gas flow operation maps of jet nozzle and turbines play important role. The published experimentally obtained dependences of the values of coefficient of discharge μ_C, velocity coefficient φ_C, total pressure recovery coefficient σ_C, coefficient of jet contraction and non-uniformity of flow in the jet nozzle "throttle" f_C, and pressure ratio π^*_{CP} are known. In this work (3D- flow simulation in conical nozzles with use of KosmosFlowWorks CAE-system was conducted by Sverchkov P.V.), nozzle gas flow pictures were analyzed using 3D-simulation in KOSMOS Flow Works CAE-system (Fig. 3). This made it possible to reveal the special features of the flow of such dependences, possibility of their representation in dimensionless parameterized form. In the same form the empirical flow characteristic of turbine nozzle apparatus was obtained, according to results of 3D calculation. Although in practice, it can be obtained also according to results of tests. Also existing data can be used, given for cone nozzles with appropriate slope angle α.

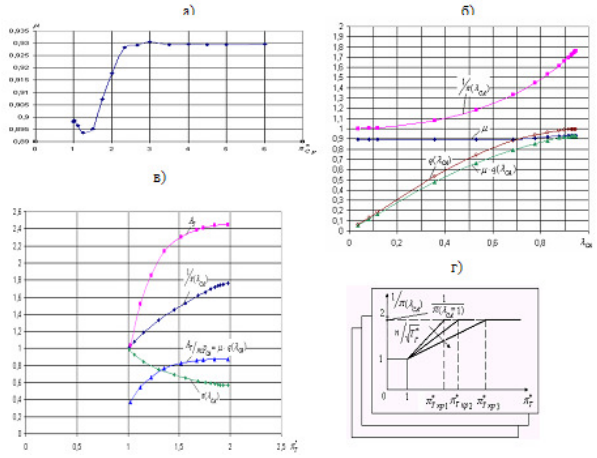


Figure 3. Dependences for Obtaining Operation Maps (Based on Example of Turbine Operation Map)

As shown by G.N. Abramovich, (1)

$$\mu = \sigma_{\Gamma} \cdot f \qquad (1)$$

From the other side, (2) and (3).

$$\mu \cdot q(\lambda_{CA}) = \frac{A_T}{mF_{CA}} \qquad (2)$$

$$A_T = \frac{G\sqrt{T_1^*}}{P_1^*} = m \cdot \mu \cdot F_{CA} \cdot q(\lambda_{CA}) \qquad (3)$$

In the work the dependence of these parameters on the turbine pressure ratio is revealed. It is taken into account, that with the approximation to the second critical operation mode (increase in the pressure ratio in the turbine), (4).

$$\lambda_{CA} \to 1,\ f \to 1,\ \mu \to \sigma_{\Gamma} \qquad (4)$$

For determining the missing turbine operation map during identification, the following procedure is proposed. If square of the nozzle cascade "thrttle" is known for the turbine, and at least one operation point is known, then ratio (5) is determined, (6) is found further, and then (7).

$$\frac{A_T}{mF_{CA}} \qquad (5)$$

$$\lambda_{CA} \qquad (6)$$

$$\frac{1}{\pi(\lambda_{CA})} \qquad (7)$$

This makes it possible to build the proposed piecewise-linear approximation (8), and to obtain the necessary dependence (9) as (10).

$$\frac{1}{\pi(\lambda_{CA})} = f(\pi_T^*) \qquad (8)$$

$$A_T = \frac{G\sqrt{T_1^*}}{P_1^*} = m \cdot \mu \cdot F_{CA} \cdot q(\lambda_{CA}) \qquad (9)$$

$$A_T = \begin{cases} \mu_{CA} \cdot m \cdot F_{CA}, \quad npu\ \pi_T^* > \pi_{TKP}^*; \\ \mu_{CA} \cdot m \cdot F_{CA} \cdot \left(\frac{k+1}{2}\right)^{\frac{1}{k-1}} \lambda_{\mathrm{CA}} \left(1 - \frac{k-1}{k+1}\lambda_{CA}^2\right)^{\frac{1}{k-1}}, \\ \text{где } \lambda_{CA}^2 = \\ = \frac{k+1}{k-1}\left[1 - \left(\frac{\pi_{TKP}^*}{\left(1 - \frac{k-1}{k+1}\right)^{\frac{k}{k-1}} \cdot \pi_T^*}\right)^{\frac{k-1}{k}}\right], npu\ \ 1 < \pi_T^* < \pi_{TKP}^*; \\ 0, \ \ npu\ \ \pi_T^* = 1; \end{cases} \qquad (10)$$

Thus for any known turbine operation map (for any branch of it) on a number of points it is possible to verify the proposed piecewise-linear approximation (11), and to analyze errors, which appear with its use.

$$\frac{1}{\pi(\lambda_{CA})} = f(\pi_T^*) \qquad (11)$$

Dimensioning (normalization) at unit operation maps representation is a necessary stage in proposed method. It allows to define the values of dimensioning parameter on a series of experimental measurements at various operation modes, and to build then the individual operation map for the given engine instance as the result of identification. However, engine operation maps in their traditional form not always convenient to make dimensionless. For decreasing the number of dimensioning parameters, author proposing the dependence in the form of (12) to transform into (13), and further into invariant dimensionless form (14) with single dimensioning parameter (15) or (16), etc., which is the same to (17).

$$\sigma, \varphi = f(\lambda),\ \mu = f(\pi) \qquad (12)$$

$$(1-\sigma) = \delta = f(\lambda),$$

$$(1-\varphi) = f(\lambda),$$

$$(1-\mu) = f(\pi) \qquad (13)$$

$$\overline{(1-\sigma)} = \overline{\delta} = f(\lambda),\ \overline{(1-\varphi)} = f(\lambda),$$

$$\overline{(\mu - 0{,}61)} = f(\pi) \qquad (14)$$

$$(1-\sigma)_o\big|_{\lambda=1} \qquad (15)$$

$$\delta_o\big|_{\lambda=1},\ (1-\varphi)_o\big|_{\lambda=1},\ (\mu - 0{,}61)_o\big|_{\pi=1} \qquad (16)$$

$$\sigma_o\big|_{\lambda=1},\ \varphi_o\big|_{\lambda=1},\ \mu_o\big|_{\pi=4} \qquad (17)$$

Authors via the carried out analysis obtained the standard invariant dimensionless one-parameter dependences of gas turbine engine units (combustion chamber, turbine, nozzle). The same dependences in the analytical form were obtained with the aid of Excel software, mainly polynomial dependences.

Conducting analysis of the series of the mean statistical operation maps of turbines of various engines, made it possible to propose using the mentioned piecewise -

linear approximation in coordinates (18) for obtaining the missing operation map of turbine.

$$1/\pi(\lambda_{CA}) = f(\pi_T^*) \qquad (18)$$

In such coordinates each branch, which corresponds to the specific rotation frequency (19), is depicted as the broken line, which consists of two parts.

$$n/\sqrt{T_T^*} = const \qquad (19)$$

This makes it possible to approximate each branch (19) of turbine operation map by universal dependence (20)

$$A_T = \frac{G_1 \cdot \sqrt{T_1^*}}{P_1^*} \qquad (20)$$

In accordance with the destination of simulation models, unit operation maps must be represented in a maximally broadband parameter range. Thus, for instance, compressor operation map must include the region of autorotation (21), the beginning of the rotor rotation (22), the sections of the pressure branches higher than surge line, zone with the reverse flow in the compressor (23), and even pressure branches for the reverse rotation of compressor (24).

$$(\pi_к^* < 1) \qquad (21)$$

$$(n_{np} \text{ от } 0) \qquad (22)$$

$$(G_{Bnp} < 0) \qquad (23)$$

$$(n_{np} < 0) \qquad (24)$$

The authors showed productivity of the idea of compressor operation map representation in standardized dimensionless form (25).

$$\frac{(\pi_к^* - 1)}{(\pi_{кБ}^* - 1)} = f_1(\bar{G}_{Bnp}, \bar{n}_{np});$$

$$\frac{\eta_к^*}{\eta_{кБ}^*} = f_2(\bar{G}_{Bnp}, \bar{n}_{np}); \qquad (25)$$

where $\bar{G}_{Bnp} = \frac{G_{Bnp}}{G_{BnpБ}}$, $\bar{n}_{np} = \frac{n_{np}}{n_{npБ}}$

Here with dimensioning, as a reference point "B", proposed to use a point on surge line (where efficiency is at maximum). This is a 2nd point of critical behavior of compressor, when simultaneously at compressor entrance and in the "throttle" area of blades the speed of sound is reached.

This approach, in contrast to the works, carried out in TSIAM, makes it possible to represent all known compressor operation maps (and also separate blade rotors) to the united universal representation (Fig. 4).

4. Approval of the developed methods and means

The approval of the developed methods and means is produced [8] by identification of special model of given engine instance (GTP-10/95 - the gas-turbine drive of the power plant GTE-10/95, based on R95SH air engine). The results of tests were used, given by research and production group “Motor”.

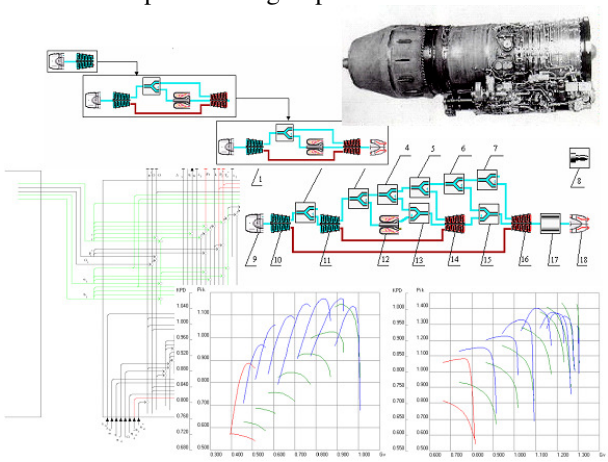


Figure 4. Sequence of Identification of Engine Operation Maps

As it often occur in practice, in this case with the tests, the parameters between the cascades (turbines and compressors) were not measured. Therefore, it is shown that during the first stage of identification it is expedient to use a conditionally single-shaft model. For this reason, an interrelation of parameters was examined for twin-shaft engine model, conditionally represented as single-shaft. In this case, in the single-shaft version, gas bleeds and injections were assigned as summary. Originally, solution search area for each parameter is determined by the physical sense (21).

$$(\mu, \varphi, \eta, \sigma < 1) \qquad (21)$$

The expected value is determined by the mean statistical parameters of given type of engine, its analogue, etc. This concerns also limitations of 2nd kind for desired operation maps and operation lines. This allowed during the first stage of identification to already obtain operation line on operation maps of combustion chamber and jet nozzle. Further, functions of change of dimensioning parameters were obtained with the use of these lines. This made it possible to obtain the individual operation maps of compressors for the given instance of gas generator GTP-10/95. Although this solution is not the only solution, due to lack of experimental data. A sequential change of model structure during the process of identification is shown in Fig. 4. Fig. 5 makes it visible that calculated values of dimensioning parameters are drifting (having a trend) on operation modes. However, the searched “non-mode” parameters should not have a trend on modes. Therefore, it is proposed to find combinations of values for such parameters by iterative method, using the principle of greatest likelihood (where this

trend for all selected parameters is minimal). Thus, in the work, further iterative determination of dimensioning parameter values carried out, for low pressure compressor and high pressure compressor operation maps. A single, most probable solution was selected (Fig. 6).

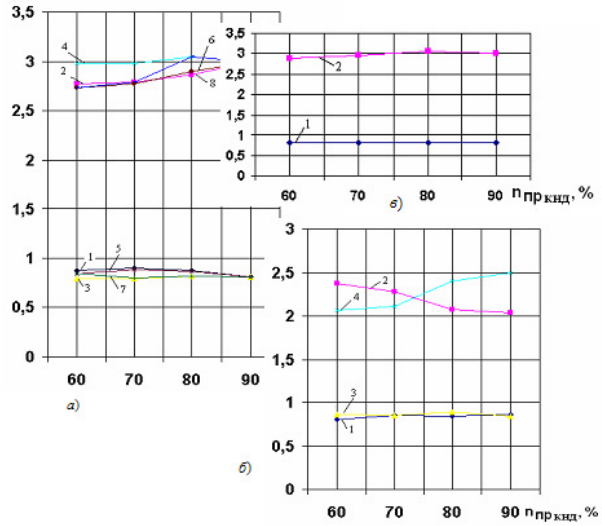


Figure 5. Results of Iterations on Determining Dimensioning Parameters for Compressor Operation Map

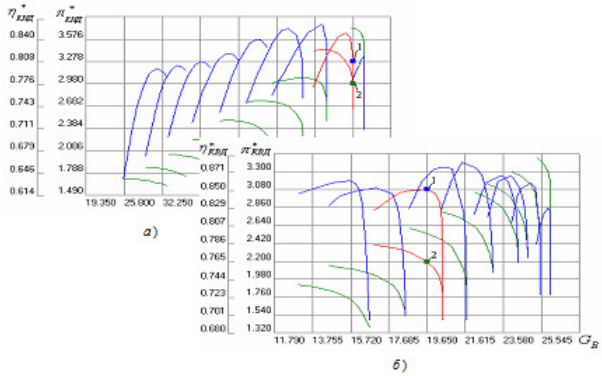


Figure 6. Identification in 3rd Iteration of an Operation Map:
A) Low Pressure Compressor;
B) High Pressure Compressor

One part of calculated parameters in this case was varied, and another was temporarily "frozen". Then the parameter groups changed their roles, using results from previous step. The parameter combination obtained, for which (22).

$$\pi^*_{КНДо},\ G_{Bпро},\ \eta^*_{КНДо},\ n_{1пр_o} = f(n_1) \approx \text{const}$$

And (22)

$$\pi^*_{КВДо},\ G_{КВДпро},\ \eta^*_{КВДо},\ n_{2пр_o} = f(n_1) \approx \text{const}.$$

This allowed to obtain most probable **individual operation maps** for low- and high pressure compressors (for the given instance of GTP-10/95). The maps were built using dimensioning parameter values, obtained as a result of the identification. It is showed that: If a priori taken operation maps are close to physical characteristics of the engine being tested, then at dimensioning process they only deformed and zoomed along X- and Y-axis.

5. Comparison of the obtained individual model to test results

Estimation of degree (and area) of adequacy for obtained individual model should be performed by comparison of calculated results and test results of the gas generator, but in different operation modes, different climatic conditions, etc. In this work, such comparison performed as with test results, and with another software calculation results as well. The alternative software was GRAD. GRAD-made identification was carried out by the workers of brigade №1 of research and production enterprise "Motor", using broad array of experimental and calculated data. Comparison results showed that the model which obtained using DVIG software and a small amount of test data, has high level of adequacy. In the case of attraction of rich experimental material owned by "Motor" and experience of their workers, the developed methods and means can give an even larger effect.

For checking the effectiveness of the developed version of the DVIG simulation system (taking into account stochastic parameters) there was carried out a simulation of GTP-10/95 gas generator, taking into account errors in input data. The estimation of determination error for the part of parameters was produced (on the basis of metrological analysis, analysis of procedures for indirect determination of values, etc.). Addition of data determination errors to the DVIG model: For input data directly, and for output data via a calculation law (as constant parameters obtained by varying input data). This allowed obtaining not only mathematical expectation, but also parameter deviation values.

6. Use of identification results for state diagnostics of a technical system

For sate diagnostics of a technical system, it is suggested to use the data about operation maps deforming during technical system operation. These data are obtained based on analysis of the trend of parameter base values (dimensioning parameters, mean statistical, individual, and obtained at the previous stages of lifecycle). With the use of simulation modeling, corresponding particular deviations for dimensioning parameters are determined for operation maps. This allows to obtain position of bounding surfaces in the space of the dimensioning parameters of operation maps. In this case, each bounding surface corresponds to particular malfunction in particular structural element, i.e., the problem of fault localization

is solved. Message about refusal (or its danger) appears, when representative point in a space of dimensioning parameters approaches a bounding surface, or intersects it. The trend itself is revealed according to results of identification of simulation models, using a measurement of monitored parameters in the process of operation. Analysis and refining of bounding surfaces is conducted based on study of the trend during operation, taking into account several different instances of a technical system. Such work is conducted for the specific arrays of instances of the given type of technical systems. Distributions of probability are revealed for particular faults in particular structure elements of a technical system. Then these standard distributions are converted into individual ones, for each particular instance of a technical system. These individual distributions are used for state diagnostics for a given instance of a technical system on various stages of its lifecycle.

Conclusion

The suggested method of composing and use of simulation models, and their consecutive identification allows to increase efficiency of computer aided lifecycle support. This is particularly true for air engines at stages of design, refining, and operation – for state diagnostics. This also corresponds to the concept of common information space.

Acknowledgement

The research is supported by grant of President of Russian Federation № 65497.2010.9 for Leading Scientific Schools.

References

1. Automated systems of design of aircraft engines. / I.A. Krivosheev, D.A. Akhmedzyanov. – Ufa: UGATU, 2002. – 61p.
2. Akhmedzyanov D.A., Gorunov I.M., Krivosheev I.A., Gumerov H.S., Kozhinov D.G., Ivanova O.N. Thermo-gas-dynamic analysis of working processes in GTE using DVIGw computer environment. The teaching aid. Ufa State Aviation Technical University. – Ufa, 2003. – 162p.
3. Akhmedzyanov D.A., Krivosheev I.A., Ivanova O.N. Dynamics of development and use of mathematical models in various development stages of GTE // Izvestiya vuzov. Aviatsionnaya tekhnika. 2003. № 3, P. 71-73.
4. Akhmedzyanov D.A., Gorunov I.M., Krivosheev I.A., Gumerov H.S., Kozhinov D.G., Ivanova O.N. "System of thermo-gas-dynamic simulation of gas turbine engines (DVIGw)" // Computer program № 2004610624 from 04.03.2004 ROSPATENT, Moscow, 2004.
5. Krivosheev I.A., Ivanova O.N. "System for simulation of gas turbine engines for solution of design-and-finishing problems, taking into account data errors (DVIGws)" // Computer program № 2004610473 from 18.02.2004 ROSPATENT, Moscow, 2004.
6. Krivosheev I.A., Kozhinov D.G., Ivanova O.N. Solution of identification, design and control problems using net simulation models // "Solving an identification, design and control tasks using network simulation models" 4-th International conference: "Inverse problems: Identification, Design and Control". July 2-July 6, 2003 Boat cruise Moscow-Kostroma-Moscow by Volga river, RUSSIA
7. Krivosheev I.A., Kozhinov D.G., Gorunov I.M., Akhmedzyanov D.A., Kharitonov V.F., Ivanova O.N. Methods and means of system development of complex objects based on simulation modeling and METACAD technology // Addition to “Information technologies” magazine №4. 2005. P. 1-31.
8. Krivosheev I.A., Ivanova O.N., Gorunov I.M. Use of simulation modeling of gas turbine engines in the stage of models identification according to tests results // Vestnik UGATU. V.6. № 1 (12), 2005. P 65-75.

Model-Driven SPD Processing and Studies of Bulk Nanostructured Materials

I.V. Alexandrov
Ufa State Aviation Technical University,
Ufa, Russia
e-mail: iva@mail.rb.ru

R.Z. Valiev
Director of Institute of Physics of Advanced Materials, Ufa State Aviation Technical University
Ufa, Russia
e-mail: rzvaliev@mail.rb.ru

V.D. Sitdikov
Ufa State Aviation Technical University,
Ufa, Russia
e-mail: svil@mail.rb.ru

Abstract[1]

Present paper demonstrates recent results of experimental and theoretical studies carried out by the authors and their colleagues with the focus on development of severe plastic deformation (SPD) techniques for producing bulk nanostructured materials, analysis of microstructure evolution leading to grain refinement, formation of bulk nanostructured states as a result of SPD, study of texture evolution to estimate the contribution of different systems of crystallographic sliding and twinning to provide material flow of the billets during SPD, analysis of activity of different potential straining mechanisms depending on the microstructure character formed as a result of SPD.

1. Introduction

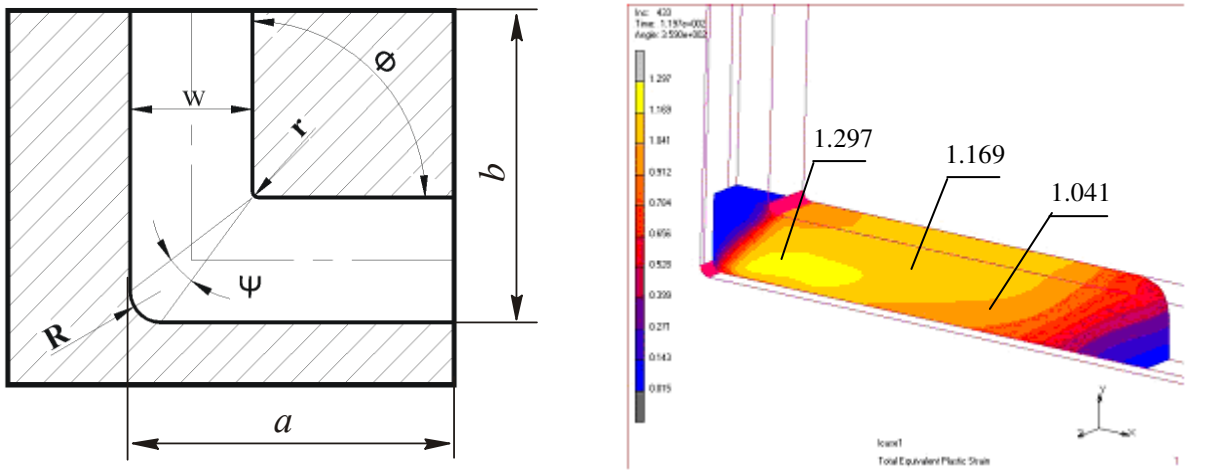
Recent years have been characterized by the birth and rapid development of new scientific trend dealing with high strained nanotechnologies and aimed at producing of bulk nanostructured materials via SPD. These metallic materials are referred to a new class of nanomaterials, possess a set of attractive physical and mechanical properties and are of a great scientific as well as an engineering interest [1,2]. This paper reports the results of the carried out research on developing the equal-channel angular pressing (the ECAP) techniques as well as processing and analysis of the obtained bulk nanostructured materials, achieved by the approaches based on computer modeling results. Special attention has been paid to the development of a multi-scale modeling, which enables disclosing the mechanisms of processing and structuring of SPD nanomaterials as well as the peculiarities of their features.

2. Development of SPD techniques

The ECAP is one of the main techniques to fabricate the bulk nanostructured billets out of various metallic materials. [2]. In the course of the ECAP a rod-shaped billet is pressed repeatedly through a die, constrained within two intersectional channels with equal cross-sectional dimensions (Fig. 1a). In spite of the apparent simplicity the processing of homogeneous nanostructured bulk billets is a highly complicated procedure, which requires an accurate design and a careful manufacturing of a die-set, choice of the material for its construction, proper selection of temperature, strain rate, lubricant and etc.
Computer simulation realized by means of finite element and variation-difference methods [3,4] has shown that a deformation zone in the billet during the ECAP is attributed to the plane of intersections of channels. This lentil-shaped deformation zone is moved along the longitudinal axis of the billet while it is pressed through the die-set. Herewith, the billet is deformed throughout the whole volume (Fig. 1b). From the performed analysis there were established the peculiarities of changes in properties of deformation zones in respect to the die-set geometry (the angle of the channel intersection, the channel diameter, the radius of channel conjunction), the number of passes of the billet through the die-set, the processing routes (orientation of the billet during successive passes), the temperature of the die-set and the billet, the lubricant, the back pressure and so on. It has been shown that the increase in the number of passes and the applied back

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

pressure, and the decrease in the level of friction result in necking of deformation zone and its accumulation at



(a) (b)

Figure 1. A schematic illustration of the ECAP die-set (a) and equivalent plastic strain distribution ε_{eqv}^{pl} along the billet length in its central plane for the 3D model (b). Cu. The 1st ECAP pass. The ambient temperature. The friction coefficient f = 0.0. (Φ = 90°, R = 1.0 mm; r = 1.0 mm, w = 8 mm, a = 43 mm, b = 68 mm).

the plane of intersection of the channels. In connection with this, the impact of the shear strain increases, the material flow becomes more homogeneous, which affects favorably the homogeneity of the microstructure and the properties of the processed billets.

The applied approach enables to optimize the stress-strain fields and the deformations, accumulated in the billets, due to variation of the above mentioned parameters. This approach allowed constructing new die-sets for the ECAP processing, including the ECAP in parallel channels and the ECAP-conform method [5]. As a result the range of metallic materials processed by the SPD techniques in order to obtain the bulk nanostructured materials has been extended. It can be applied to both ductile metallic materials and hard-to-deform low-ductile metals and alloys. The processing finalizes in enhanced homogeneity of the produced materials increase, size of the obtained billets, metal yield and productivity. By the moment the long-sized bulk nanostructured billets can reach up to a few meters in length with a few millimeters in diameter. The production of sheet nanostructured materials has also been started here. These facts point out the perspectives of commercial production of bulk nanostructured billets.

3. Analysis of microstructure and crystallographic texture

During the ECAP processing a special attention is given to the aimed development of homogeneous bulk nanostructured states with the controlled parameters of a microstructure (the volume fraction of grain-boundaries, mainly high-angle grain boundaries, degree of nonequilibrium in grain boundaries, crystallographic texture and so on). From the point of view of the maximum efficiency of the computer modeling, applied for prediction and analysis of the obtained experimental values of the above mentioned metallic material parameters, subjected to the SPD technique, now it is possible to determine the development process modeling of the crystallographic textures.

The analysis performed mainly by a visco-plastic self-consistent model has shown that the simple shear texture is the most typical for the texture components, formed during the ECAP [6]. At the same time, heterogeneity of the material flow in the processed billets results in activation of stretching and tension strain components, which can affect significantly the character of the appearing crystallographic orientations. Computer modeling of the texture development processes has allowed estimating the activity of various deformation modes and the activity of various slip and twinning systems during SPD and, thus, to the peculiarities of the deformed material's flow comprising its homogeneity.

From this point of view the most essential are the results of computer modeling, applied to the analysis of texture development, according to the number of passes and routes of the ECAP. The carried out analysis has shown that the grain size of the processed metals can affect the activity of various systems of sliding and twinning.

Microstructure refinement occurs in the process of the SPD. In this regard the experiment on Ti with the HCP crystallites showed the activity of crystallographic twinning during the first pass of the ECAP and its diminishing during the following passes [7]. Thereby, the activity of various sliding and twinning systems is significantly affected by the number of passes (Fig. 2). Further microstructure refinement with the extremely high strain degrees, achieved by the SPD, results in a

scattered crystallographic texture, which testifies to activation of peculiar modes of the plastic deformation

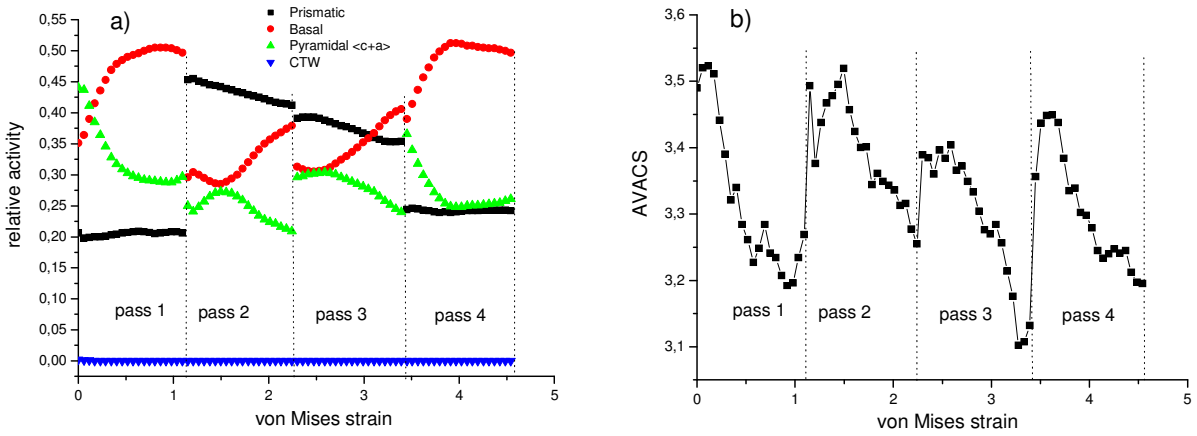


Figure 2. Relative activity of sliding and twinning systems (a) and an average system activity per grain (b) as a function of strain (Ti, ECAP, T = 723 K). Route B_C.

[8]. From this point, application of computer modeling to analysis of such processes is highly promising. To describe the processes, considering the effect of a structure refinement, a simulation scheme, which uses the visco-plastic self consistent model coupled with the disclination criterion for the grain subdivision, has been worked out [9]. It has shown that the configuration of the defects, which appear due to the dislocation activity, makes energetically favorable the grain subdivision by low-angle grain boundaries, which accumulate misorientations with further straining and it results in an intensive structure refinement. The optimal agreement with the experimental data has been achieved by the multi-scale modeling when the visco-plastic self consistent model accounted both the consideration of grain refinement and the 3D FEM results, taking into account the material flow at a macrolevel [10,11].

4. Analysis of deformation mechanisms

Microstructure with ultra small grain size results in the development of peculiar deformation mechanisms in bulk nanostructured materials produced by the SPD. It is specifically expressed in the absence of strain hardening at high strength. The analysis of dislocation sources and sinks both in the grain interior and at the grain boundaries, performed by dislocation kinetic modeling on Cu samples, subjected to various number of ECAP passes, has shown the following. After 8 ECAP passes through the route B_C the role of annihilation processes, conditioned by capture and non-conservative motion of edge dislocations, has been reduced. The mechanism of dislocation generation has been suppressed on the forest of dislocations.

The mentioned changes reflect qualitative reconstructions in the structure of fragment boundaries. The boundaries become practically impermeable and the fraction of non-excess sessile dislocations is negligibly small. Insignificant strain hardening of the materials subjected to ECAP is explained by the fact that growth of the activity of the sources is compensated by the enhancement of annihilation processes [12].

4. Conclusion

The achieved results presented in the paper testify to the importance of computer modeling in the study of SPD processing. The developed approach of multi-scale modeling made it possible to achieve new data on regularity of microstructure and crystallographic texture formation as well as deformation behavior of nanostructured SPD materials.

References

1. Valiev R.Z., Islamgaliev R.K., Alexandrov I.V. “Bulk nanostructured materials from severe plastic deformation”. *Progr. Mater. Sci.* 2000; 45: 103-189.
2. Valiev R.Z., Estrin Y., Horita Z., Langdon T.G., Zehetbauer M.J., Zhu Y.T. “Producing bulk ultrafine-grained materials by severe plastic deformation”. *JOM*. 2006; 58: 33-39.
3. Zhernakov V.S., Budilov I.N., Raab G.I., Alexandrov I.V., Valiev R.Z. “A numerical modeling and investigations of flow stress and grain refinement during equal-channel angular pressing”. *Scr. Mater.* 2001; 44: 1765-1769.
4. Alexandrov I.V., Budilov I.N., Krallics G., Kim H.-S., Yoon S.Ch., Smolyakov A.A., Korshunov A.I., Solovyev V.P. “Simulation of equal channel angular pressing”. *Mater. Sci. Forum*. 2005; 503-504: 201-206.
5. Raab G.I. Dissertation for Dr. Sci. degree, 2009.
6. Alexandrov I.V., Zhilina M.V., Scherbakov A.V., Bonarski J.T. “Multi-scale investigations of severe plastic deformation”. *Mater. Sci. Eng.* 2005;

A 410-411: 332-336.
7. Alexandrov I.V., Sitdikov V.D., Bonarski J.T. “The evolution of crystallographic texture in CP Ti, subjected to equal-channel angular pressing”. *Vestnik USATU*. 2009; 12: 76-82.
8. Enikeev N., Schafler E., Zehetbauer M., Alexandrov I., Valiev R. “Observations of texture in large scale HPT-processed Cu”. *Mater. Sci. Forum.* 2008; 584–586: 367-374.
9. Nazarov A., Enikeev N., Romanov A., et al. „Analysis of substructure evolution during simple shear of polycrystals by means of a combined viscoplastic self-consistent and disclination modeling approach“. *Acta Mater.* 2006; 54: 985-995.
10. Enikeev N.A., Abdullin M.F., Nazarov A.A., Beyerlein I.J. **“**Modeling grain refinement in FCC metals during equal-channel angular pressing by route C”. *Int. J. Mater. Res.* 2007; 98: 167-171.
11. Smolyakov A.A., Solovyev V.P., Korshunov A.I., Enikeev N.A. ”Three-dimensional numerical simulations of multi-pass equal-channel angular pressing by a variation difference method and comparison with experiment”. *Mater. Sci. Eng.* 2008; A 493: 148-159.
12. Alexandrov I.V., Chembarisova R.G. ”Development and application of the dislocation model for analysis of the microstructure evolution and deformation behavior of metals subjected to severe plastic deformation”. *Rev. Adv. Mater. Sci.* 2007; 16: 51-72.

Analysis of Unique Properties of Nanostructured Alloys Produced by Severe Plastic Deformation

N.A. Enikeev
Institute of Physics of Advanced Materials
Ufa State Aviation Technical University
Ufa, Russia
e-mail: carabus@mail.rb.ru

M.Yu. Murashkin
Institute of Physics of Advanced Materials
Ufa State Aviation Technical University
Ufa, Russia
e-mail: maxmur@mail.rb.ru

I.V. Alexandrov
Physics Department
Ufa State Aviation Technical University
Ufa, Russia
e-mail: iva@mail.rb.ru

R.Z. Valiev, Dr. Sc., Professor
Institute of Physics of Advanced Materials
Ufa State Aviation Technical University
Ufa, Russia
e-mail: rzvaliev@mail.rb.ru

Abstract[1]

It is shown, that ultrafine-grained (UFG) alloys produced by severe plastic deformation (SPD) can be characterized by a combination of unique properties.
As an example, UFG Al alloys are found to exhibit superstrength, considerably exceeding the Hall-Petch predictions for the ultrafine grains. The phenomenon is explained by the unique combination of ultrafine-grained structure and deformation-induced grain boundary segregations, which may affect the emission of dislocations from grain boundaries and their mobility.
Besides, nanostructuring of materials can lead to their enhanced radiation stability. It can be explained by the fact due to nanostructured materials have considerably increased volume fraction of grain boundaries. The grain boundaries can actively absorb dislocations and function as effective drains for point defects and, as a result, increased radiation resistance. The first experimental evidences of such phenomena are reviewed.

1. Introduction

Nowadays main principal possibilities for elaboration of new materials with the required properties at expense of their chemical composition variation have been practically exhausted. In the given circumstances the special importance is assumed by drastic improvement of the properties of already existing materials through modification of their microstructure parameters at expense of new approaches to their treatment.
Nanostructuring is the new and promising way to enhance properties of metallic materials for advanced functional application. To date, it is well established that bulk nanostructured metallic materials can be produced successfully via microstructural refinement using SPD techniques [1]. SPD processing is promising for many potential applications, as it allows to significantly improve functional properties of commonly used metals and alloys [1-2].
The given paper reports on abnormally high strength and elevated radiation tolerance in alloys nanostructured by SPD.

2. Superior strength of SPD Al alloys

Commercial alloys 1570(Al – 5.7%Mg – 0.32%Sc – 0.4%Mn, wt.%) and 7475(Al – 5.7%Zn – 2.2%Mg – 1.6%Cu – 0.25%Cr, wt. %) were nanostructured via high pressure torsion (HPT) – the most effective SPD technique for structure refinement at expense of applying enormous strains under high applied pressure. The applied pressure of P=6 GPa with number of rotations n=20 were used in the samples processing at room temperature. The obtained UFG samples were in form of discs with a diameter of 20 mm and 0.8 mm in thickness which is suitable for tensile strength testing [3]. The obtained samples were studied with the help of transmission electron microscopy (TEM), X-ray diffraction (XRD) to reveal structure parameters (grain size, lattice parameter). Tensile tests at room temperature at a strain rate of 10-4s-1 were performed to measure mechanical properties.

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

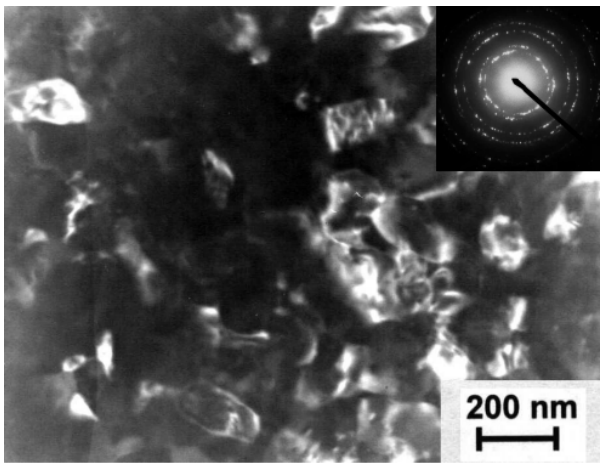


Figure 1. A dark field TEM image of the UFG 1570 alloy processed by HPT

TEM analysis proved that the HPT processing of the alloy contributed to complete refinement of the initial coarse-grained (CG) structure into a UFG one. Homogeneous structures with a mean grain size of about 100 nm were formed in HPT processed alloys as shown on Fig. 1 for the example of 1570 alloy [3]. The SAED image (Fig. 1) represents numerous multiple spots uniformly distributed along circles. It testifies that the microstructure consists of mainly high angle grain boundaries.

It was also determined that HPT-processing has a visible effect on the value of crystal lattice parameter *a* of Al alloys - in the alloy 1570 it reduced considerably in comparison with the initial state – from 4.0765±0.0001Å to 4.0692±0.0003Å. It is known, decreasing of *a* is linked to decreasing of Mg in solid solution Al-Mg [4]. Thus, the observed change of Δa=0.073 Å after HPT is equivalent to loss of ~1.6 wt. % Mg in solid solution of the 1570 alloy. One can suppose that the alloying elements were deformation-driven from solid solution to grain boundary area. This suggestion is confirmed by latest 3D atom probe analysis results, evidencing formation of Mg segregations in the grain boundary region after HPT [5]. The revealed peculiarities in the alloys' microstructure well correlate with their mechanical properties.

Table 1. Mechanical properties of 1570 and 7475 alloys in CG and UFG states.

Alloy	1570 CG	1570 UFG	7475 CG (T6*)	7475 UFG
σ_{02}	190	905	460	920
UTS	376	950	525	980

* - conventional heat treatment, providing maximal strength in Al alloys [4]

Table 1 provides results of the mechanical tests for 1570 and 7475 alloys. The exhibited strength values exceed those of conventional high-strength heat-treated Al alloys and a number of UFG Al alloys produced earlier [6,7]. The strength of the UFG 1570 alloy is even higher than it follows from the Hall-Petch dependence (Fig. 2), which indicates the presence of additional factors leading to alloy hardening .

Such unusual change in the mechanical properties of the UFG alloy may be associated with the abovementioned segregations of solute elements along the boundaries of formed UFG grains. Changes of the phase composition in grain boundary area can affect considerably the activation volume and, correspondingly, the stress, required for emission of dislocations from the grain boundaries [8].

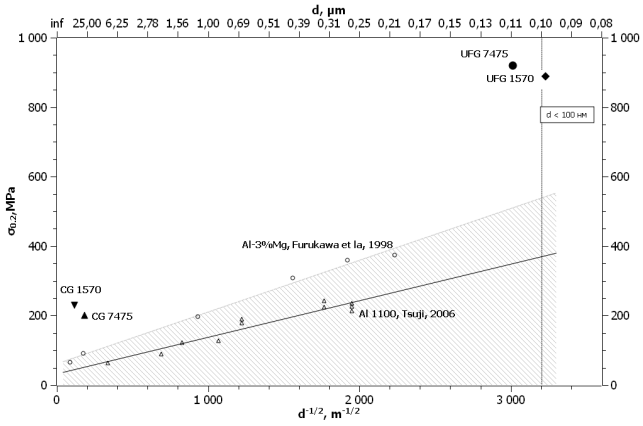


Figure 2. The Hall-Petch relation for the alloys 1100 [9], Al-3%Mg [10] and data on the yield stresses of the UFG alloys 1570 and 7475

Thus, extraordinary high strength, considerably exceeding the Hall-Patch predictions for the given grain size, observed in HPT processed UFG Al alloys, can be attributed to not only the UFG structure of the material, but also to formation of segregations of alloying elements in grain boundary area.

3. Radiation tolerance of nanostructured alloys

One of the most actual problems in the field of energetic engineering is the development of new unique materials with increased radiation tolerance [11]. Critical components of fusion/fission reactors are to be characterized by high long-range reliability at severe radiation, elevated temperatures and aggressive environment. Properties of solids are subjected to degradation being exposed to intensive high-energy particles flows (fast neutrons) and radiation (γ and X-rays) at long-term usage. It leads to changes in the materials' microstructure, which are expressed in radiation induced crystallite lattice damages (vacancies and their clusters, interstitial atoms, pores and dislocation loops) and nucleation of radiation-induced phases. It entails decreasing the materials functional properties, critical to reliable operation of nuclear reactors' internal structures and may cause their

malfunction. Development of radiation-resistant materials for energy applications is a topical task for modern materials science with relation to both versatility together with security of the leading-edge energy technologies and ecological security up to risks reductions in conditions of climate fluctuation and man-made disasters. The problem of prime importance is development of new functional materials characterized by high radiation resistance to be used in new generation nuclear plants with increased fuel cycles' duration.

Nanocrystalline metals and alloys are attractive as advanced interface-controlled structural materials due to the expectation of increased irradiation resistance. Molecular dynamics simulations demonstrated that the grain boundary structure at the nanoscale regime strongly affects the primary damage state, with the grain boundary acting as a sink for interstitials [12,13]. In ion irradiated nanocrystalline Pd and ZrO2, for example, the damage, expressed in density of point defect clusters observed by TEM, decreased by a factor of 3 to 4 when the grain size was reduced from 100 nm to 40 nm [14]. These results are a clear indication that nanocrystalline materials can exhibit enhanced irradiation resistance.

SPD methods enable to form nanostructure states in notably bulk billets [1]. Nanostructuring by SPD processing rebounds for radiation-induced defects to an opportunity to sink into an extensive network of nonequilibrium high-angle grain boundaries of BNM. As a result, the radiation resistance of BNM may be much higher than that of coarse-grained analogues. The combination of reported extraordinary mechanical properties and increased radiation tolerance could make SPD materials to be of a high potential being applied to the energy technologies.

Latest results show that SPD processed materials indeed exhibit elevated resistance to electron and ion radiation exposure [15,16]. It is expressed not only in concentration reduction of radiation-induced defects in steels [15], but also in preserving of high degree of order in crystal structure of irradiated ordered alloys [16].

Up to the present day BNM radiation resistance investigations have been just started, but have already demonstrated promising results and important tendency in enhancing a radiation tolerance of metals by their nanostructuring. The origin and physical background of radiation-resistance behaviour of BNM is a primary task to be investigated for establishing a scientific fundament to improve essentially radiation-resistant properties of commercial alloys used in energy industry.

4. Conclusion

The presented results demonstrate that the nanostructured alloys processed by SPD possess unique combination of functional properties to be of high potential for advanced applications.

References

1. Valiev R.Z., Islamgaliev R.K., Alexandrov I.V. "Bulk nanostructured materials from severe plastic deformation" *Progr. Mater. Sci.*, 2000, 45:103-189.
2. Valiev R.Z., Estrin Y., Horita Z., Langdon T.G., Zehetbauer M.J., Zhu Y.T. "Producing bulk ultrafine grained materials by severe plastic deformation", *JOM*, 2006, 58:33-39.
3. Murashkin M., Kilmametov A., Valiev R. "Structure and Mechanical Properties of an Aluminum Alloy 1570 Subjected to Severe Plastic Deformation by High-Pressure Torsion", *The Phys. Met. Metallogr.*, 2008, 106:90-96.
4. Hatch J. E. (Ed.) "Aluminum: Properties and Physical Metallurgy", ASM Intlnl, 1984.
5. Valiev R.Z., Enikeev N.A., Murashkin M.Yu., Kazykhanov V.U., Sauvage X. submitted to Scr. Mater 2010.
6. Markushev M.V., Murashkin M.Yu. "Structure and Mechanical Properties of Commercial Al-Mg 1560 Alloy after Equal-Channel Angular Extrusion and Annealing", *Mater. Sci. Eng. A*, 2004, 367:234-242.
7. Roven H. J., Nesboe H., Werenskiold J.C., Seibert T. "Mechanical properties of aluminium alloys processed by SPD: Comparison of different alloy systems and possible product areas", *Mater. Sci. Eng.*, 2005, A410–411:426-429.
8. Valiev R.Z., Enikeev N.A., Murashkin M.Yu., Aleksandrov S.E., Goldshtein R.V. "Ultrastrong Ultrafine-Grained Aluminum Alloys Produced by Severe Plastic Deformation", *Doklady Physics*, 2010, 55:267
9. Tsuji N. "Unique Deformation Behaviors of the Ultrafine Grained Aluminum Alloys Fabricated by Accumulative Roll Bonding". In: Zhu Y.T., Varyukhin V. (eds.), *Nanostructured Materials by High-Pressure Severe Plastic Deformation.* Springer Netherlands, 2006, pp. 227-234.
10. Furukawa M., Horita Z., Nemoto M., Valiev R.Z., Langdon T.G. "Factors influencing the flow and hardness of materials with ultrafine grain sizes", *Phil. Mag. A.,* 1998, 78:203-215.
11. Rubel M. "Reactor aspects of fusion: issues related to materials, radioactivity and radiation induced defects", *Trans. Fusion Sci. Tech,* 2006, 49:465-473.
12. Samaras M., Derlet P. M., Swygenhoven H. V., Victoria M. "Atomic scale modelling of theprimary damage state of irradiated fcc and bcc nanocrystalline metals", *J. Nucl. Mater.*, 2006, 351:47-55.
13. Voegeli, W. Albe, K. Hahn, H. "Simulation of grain growth in nanocrystalline nickelinduced by ion irradiation", *Nucl. Instr. Meth. B,* 2003, 202:230-235.

14. Rose M., Balogh A. G., Hahn H. “Instability of irradiation induced defects in nanostructured materials”, Nucl. Instrum. *Methods Phys. Res. Sect. B,* 1997, 127-128:119-122.
15. Kilmametov A.R., Gunderov D.V., Valiev R.Z., Balogh A.G., Hahn H. “Enhanced ion irradiation resistance of bulk nanocrystalline TiNi alloy”, *Scripta Mater.*, 2008, 59:1027-1030.
16. Radiguet B., Etienne A., Pareige P., Sauvage X., Valiev R. “Irradiation behavior of nanostructured 316 austenitic stainless steel”, *J. Mater. Sci.*, 2008, 43:7338-7843.

The information technologies in interior design

Sandu Olga
Department "Advertising and design"
Izhevsk state technical university
Izhevsk, Russia
e-mail: «solimy@rambler.ru»

Abstract[1]

The review of possibilities of using IT in interior design was represented to determine a problem of design automation of decorative finishing.

Introduction

The expectations of publishing: we shall consider the possibilities of using IT in interior design, the basic aspects of design and the structure of software meant for designing of interior mouldings.

Publishing

The information technologies serve to designer not only as universal instruments of artistic perfecting of new products, but also as possibilities of copying of goods. The potential of IT in design is very broad, but not boundless. A point of time has come, when designer wants to increase the possibilities of IT by adding new software. It`s clear that it is impossible to reach such purpose without knowledge in spheres of modern programming technologies and theory of design. We shall consider the using of software in general process of designing and in field of interior design.

The activity of design is one of the most difficult. The complexity is determined by versatility. Design has spread on almost all cultural spheres and designer combines in itself an artist and a technologist. At present, designer skills are applied both in artistic shaping and in activity arrangement. Besides, to create subjective ambience designer gets thorough all stages of producing of goods, beginning from concept formation and ending by model engineering of produce. The using of IT has improved activity of design greatly.

There are some stages of design [1]: exploratory design, designing itself and management of technological process. The automation of stages of design undertakes the decision only a part of problems.

The exploratory design is one of the most difficult and uncontrolled stage of designing. On this stage the creative work of project is carried on and trend of development is shaping. It is that level on which the creative methods are used and it`s necessary to increase the time of concept formation. Such this work is performed by designer. The detail engineering and the management of technological process are lend itself a set of standard operations. They are connected with using of ideas, received after exploratory design. It is more effectively to combine a human and automated designing on this stages. The automatization of engineering enables to reduce duration of product development and allows to redistribute time between steps of design activity.

To choose the subject of automatization, we shall select the main aspects of designing: *stylistic* and *functional solution*, *composition* and *manufacturing capabilities* (figure 1). The interior design is based on synthesis of aesthetical and pragmatic decisions, directed on improvement of life quality of person. Often it is artistic presentation that creates the comfortable and holistic image of interior. We shall consider these aspects in designing of building shell and decorative finishing of interior.

The automation is possible both in step of constructing interior itself and in decorative finishing. But not all aspects of design are pertaining to automation. Styling and planning of functions are accomplished by designer on the step of exploratory design. At present, the information technologies have been used basically to decide application tasks, for example, compositional and technological tasks. So there is a broad spectrum of software in arsenal of designer. It is used for constructing building shell, creating working documentation, 3d modeling, texture mapping, visualizing and locating the objects in interior. In spite of abundance of such sort of software, it`s necessary to make specific program product. The most perspective in given attitude becomes the software, explicated in field of designing of decorative finishing.

There are a lot of methods of interior finishing. Amongst them a decorative moldings remain one of the most actual facilities of stylistic decision. They aptly execute aesthetic problems, expressing the main idea of interior and creating holistic image. With functional standpoint decorative molding allows to shape the space of interior correctly, to highlight key points or hide the defects. For instance, moldings help to zone the ambiences of interior

[1] **Proceedings of the International Workshop "Innovation Information Technology – Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

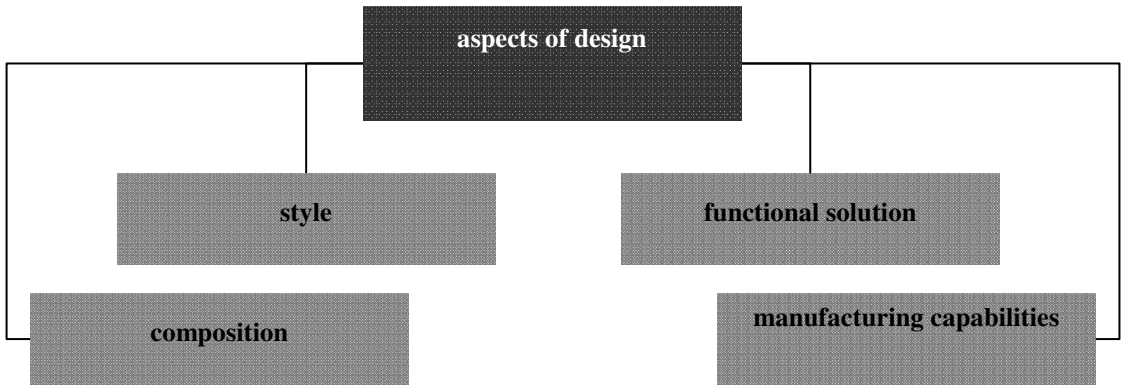


Figure 1. The aspects of design

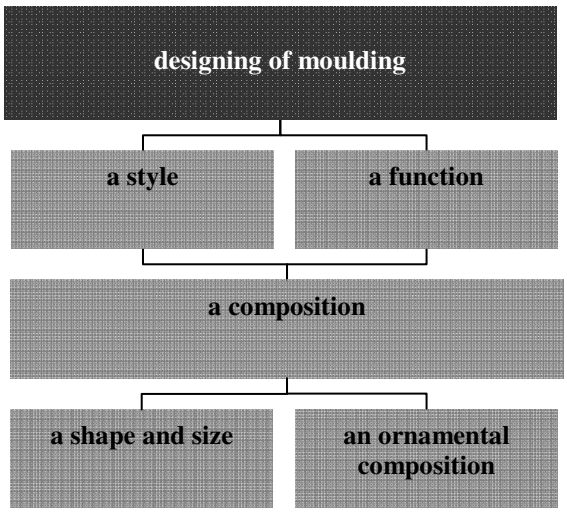


Figure 2. Scheme of software for designing of moulding

The software improving decoration can be a part of integrated system used in interior design. The structure of software consists of modules (figure 2): a "*style*", a "*function*", a "*composition*", a "*shape and size*" and an "*ornamental composition*". The style is assigned by designer on the stage of exploratory design. In module "function" one can choose the type of onlay. The function of moulding depends on structure of interior.

The process of decorative finishing consists of some steps: selecting the way of finishing, shaping and locating of decorative elements in interior. As a rule, the decor is drafted manually and vectorized by designer. Such activity takes a lot of time. It is labor-intensive process that we need to improve. So, it is actually to automate the process of designing of décor.

The interior ambience is a combination of building shell (ceiling, walls, floor) and items of apartment (furniture and equipment). So, we shall select the functional types of mouldings: *wall-type moulding*, *overhead cornice*, *arched moulding, furniture onlay* and *specific mouldings* (for interior equipment). In modules "composition of decor" and "ornamental composition" one can choose a positional option of onlay. To select the types of compositions we shall consider following concepts: a category of rhythm and a theory of symmetry (tabl.1). So, we have the basic compositional types: "rosace", "frame", "lineal", "grid" and "combined picture". Depending on number of axes, the composition can be *single-layered* and *multilayered*. The shape and size of onlay one can select in module "*shape and size*".

Table 1 Types of compositions

Rhythm	Type of symmetry [2,3]	Types of compositions
Meter	Radial	"Rosace"
	Translational	"Lineal", "Frame"
	Meshy	"Grid"
Rhythm	Combined	"Combined picture"

Conclusion

Considered software will allow to design the decorative moulding. One can choose it`s forms, sizes and the variants of ornamental composition. The software will allow to see the onlay in three-dimensional view and to contribute necessary changes on step of engineering.

References

1. “Design. Illustrated dictionary”. Moscow, Russia, 2004.
2. Weyl H. “Symmetry”. Moscow, Russia, 1968.
3. Sandu O. “Compositional features in ornaments”. Design. Materials. Technology, Vol. 2. Saint Petersburg, Russia, 2008, pp 21-24.

An Overview of Software-based Support Tools for ISO 26262

Denis Makartetskiy, Davide Pozza, Riccardo Sisto
Dipartimento di Automatica e Informatica, Politecnico di Torino
e-mail: {denis.makartetskiy, davide.pozza, riccardo.sisto}@polito.it

Abstract

Safety in the automotive domain is becoming more and more important with the ever increasing level of complexity in emerging technologies built-in into the cars. As a stimulus for industry to refine its safety measures related to electrical, electronic and software systems in the cars, the ISO 26262 standard has been recently introduced.
Developing safety-related systems according to this standard in an efficient and effective way requires adequate computer-aided support. For this reason, some initiatives towards software-based supporting tools for ISO 26262 were recently started.
This paper gives an account of the main such initiatives after recalling the main features of the ISO 26262 standard.
In particular, we briefly discuss how the main activities from ISO 26262 such as hazard analysis and risk assessment, functional safety concept, safety validation, etc can have a software support, and what is the state-of-the-art.

1. Introduction

Nowadays, electronic components are pervasive in road vehicles. Indeed, during the last 30 years the majority of innovations on road vehicles (e.g. electronic injection, airbags, ABS, ESP) have been based on electronic embedded systems, and this trend seems confirmed to continue in the future. Therefore, functional safety (i.e. the safety that depends on the system or equipment operating correctly in response to its inputs, including the safe management of likely driver errors, hardware failures and environmental changes) is acquiring more and more importance for the electronic systems that have to be integrated inside cars.
Up to now, automotive companies have developed similar, but custom, safety-related methodologies and processes, and have used pretty much the same set of techniques (e.g. FMEA, FTA) for the development of safety-related electronic embedded systems. However, at the same time, they are not sharing a common view and consideration of the safety of the produced items. In this context, ISO 26262 should become the functional safety reference standard. Although it is not yet published in its final form (it should be in 2011), it represents the shared effort of car makers, OEMs and Tier-1 suppliers to establish a common way to understand and consider the safety concept and its importance, when designing and developing embedded systems for road vehicles. As a consequence, car-makers have already started activities directed to transform their current processes to produce systems that can be ISO 26262 compliant.
One of the problems that emerge when applying this transformation is how to organize the new processes and how to adequately support them by software tools. A first commercial tool that partially responds to this need has recently appeared. Moreover, some collaborative initiatives have emerged in the recent panorama: French car-makers have started in 2007 a project called EDONA (Environnements de Développement Ouverts aux Normes de l'Automobile), while in Italy a project called SiSMA (Sicurezza Funzionale dei Sistemi Meccatronici Automotive) is just starting.
From the technical point of view, an important shared idea about the supporting software tools is that they should provide a means for modelling at the system-level and at different abstraction levels. This kind of models should then be used as a unifying basis on which the safety-related activities are built and synchronized. This modelling facility should be at least semi-formal, in order to respond to the requirements of ISO 26262.
Among the various available possibilities, three semi-formal languages have mainly been considered: plain UML, SysML, and EAST-ADL2 [4]. While the first two are general-purpose solutions, the latter is a domain specific specialization of the former, specifically tailored for modelling automotive embedded systems, that enriches the official AUTOSAR modelling language with additional abstraction levels. AUTOSAR is another important standard that focuses on the design and development of software for automotive embedded systems.
The aim of this paper is to present an overview of the above mentioned current initiatives towards software tools supporting ISO 26262.
In section 2, general information about the ISO 26262 and AUTOSAR standards is outlined. In section 3, the existing projects aiming at bringing integrating solutions for industry to apply ISO 26262 techniques in practice are analyzed. Finally, section 4 concludes.
The definitions that are necessary for understanding the domain will be given where it is needed.

2. ISO 26262 and AUTOSAR

ISO 26262 is an upcoming standard that adapts ISO/IEC 61508 (a standard concerning the safety of systems, that is applicable to all kinds of industry) to the automotive industry. In particular, ISO 26262 addresses the safety related systems that have to be installed in series production passenger cars (with a maximum gross weight up to 3500kg), that are composed of one or more electrical or electronic (E/E) systems.

At the time of writing, ISO 26262 is in the state of a draft international standard (DIS). The figure below shows the simplified structure of its development process (the V-model).

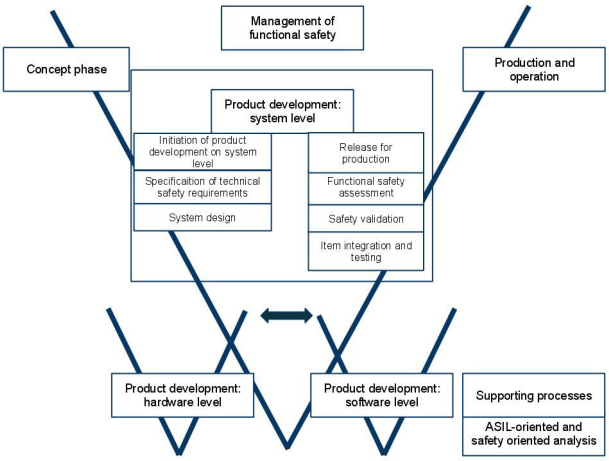


Figure 1. V-Model from ISO 26262

As it is shown in the figure, the V-model development applies to software and hardware development independently on each other and on the overall development process as well.
The standard prescribes the V-model for the development process and describes how functional safety has to be managed during the whole lifecycle of E/E safety-related systems, while providing guidance in the selection of (core and supporting) processes within product development, in function of the outcome of a specific safety analysis methodology. Indeed, the standard is centered around the concept of Automotive Safety Integrity Level (ASIL), that is a qualitative measure of the needed integrity level (i.e. the probability with which a system correctly performs its safety-related functions). The ASIL is determined by means of hazard analysis and risk assessment. According with the standard, since the beginning of the development of systems, each one of its intended functions has to be analyzed with respect to possible hazards. Hence, in function of the probability of exposure to an hazard, the possible controllability by a driver and the severity of a critical event, the risk is estimated and an ASIL is determined. Four ASIL levels have been defined, running from A (lowest) to D (highest). ASILs have to be mapped on the safety requirements that have to be generated to avoid/reduce the identified risks. Thus, it encourages making a focus on safety-critical functions, while not wasting a lot of time on non-critical ones. Therefore, the standard provides requirements for the whole lifecycle and guidance in choosing adequate methods (e.g. hazard analysis, risk assessment, and safety analysis methods) and procedures (e.g. safety, requirements, and document management) to achieve the required ASIL for the developed product.
The standard defines some new terminologies and concepts. It defines the word “item” as a system or array of systems or a function to which ISO 26262 is applied and the process of collecting and describing an item as “item definition”.
The concept phase consists of defining the preliminary architectural and functional design of future system elements and of performing their safety analysis. Safety analysis results in safety requirements that have to be satisfied in order to achieve safety goal (formalized conditions of element's safe functioning). The concept phase is followed by specification of safety requirements into low-level technical requirements and system design. On the hardware and software level the implementation of system functionalities along with technical requirements is done. Then, verification and validation counterparts for safety goal specification and system design take place. Eventually, provided that during system production, safety requirements are not violated, system goes to operation.

Although ISO 26262 derives from ISO 61508, it differs in some valuable points. While ISO 61508 mainly covers industrial equipment and process plant, which are usually produced in small numbers, ISO 26262 focuses on E/E systems for series production cars, hence the standard also covers the requirements for the production of systems in series. Moreover, it is worth noting that ISO 26262 provides much more information and guidance for qualifying and classifying software tools than ISO/IEC 61508. Furthermore, it has to be noticed that controllability was not foreseen by ISO/IEC 61508 to compute Safety Integrity Levels.
There is some connection between ISO 26262 and another important automotive standard called AUTOSAR (AUTomotive Open System Architecture). This is an ongoing motion in the automotive world, started in 2003 and directed to building common open standardized software architecture. For more information about AUTOSAR see, for example, [10], [11], [12].
Since AUTOSAR addresses also safety critical embedded software, it aims at showing relevant compliance with related safety standards (ISO 26262 in this case).

In fact, AUTOSAR defines both a software architecture and a supporting methodology to develop E/E software systems for the automotive domain but cannot guarantee functional safety of such systems by itself. Thus, the implementation of safety-related embedded systems using AUTOSAR has to be done with compliance to related safety standards designed for the automotive domain.
Starting from release 4, the AUTOSAR standard proposes some technical information required proving the ISO 26262 compliance for AUTOSAR members, but it does not provide procedures or activities that address the safety problem by itself.
The full responsibility for implementing the functional safety mechanisms described inside the AUTOSAR framework fully resides on the implementer who will have to fulfill all the specific safety related regulations. This means that AUTOSAR does not include implementation of safety activities into their shared software platform, leaving them for implementation by car makers independently on collaborative AUTOSAR movement, thus competing with each other on this implementation.
The approach to functional safety in AUTOSAR with respect to ISO 26262 concerns mainly the Safety Element out of Context (SEooC) that is a safety element for which an item does not exist at the time of the development. According to ISO 26262, during a system development process, a safety element can be developed as an item (with stricter requirements and more tediously) or as a SeooC (when requirements are substituted with assumptions). For details about SEooC, see [1].
As for modeling, the official AUTOSAR modeling language has been developed as an UML/SysML profile. For details, see [9].

3. Supporting tools and initiatives.

Using software tools can extremely facilitate some of the activities required by ISO 26262 for the development of safety-critical electronic embedded systems. Indeed it would be hard if not impossible to accomplish all the requirements of the standard without an adequate software tool support.
Among the tasks which have to be done in compliance with ISO 26262 and that should be supported by software tools there are: item definition, ASIL determination, ASIL decomposition, hazard analysis and risk assessment activities, safety goal definition and safety requirements allocation, V&V activities safety validation, configuration and change management, etc.
While for most of these activities support software tools already existed before the introduction of ISO 26262, the standard requires a unified management of safety-related activities whereby the various tools need to be integrated as a part of a single framework. This need arises, for example with respect to traceability requirements.
At the best of our knowledge, there are only two main initiatives collecting software solutions for the aforementioned activities under a single roof: the Medini Analyze software tool and the EDONA project and platform. In this chapter we will analyze how exactly they cover these activities and the requirements coming from the standard.

3.1. Medini

Medini Analyze aims at covering all the main ISO 26262 activities during the system development process with a particular focus on safety analysis. It brings together functional architecture design and functional safety analysis, making the main accent on hazard analysis and risk assessment. The structure of the work flow in Medini Analyze reflects corresponding parts of the ISO 26262 V-Model.
Medini Analyze naturally introduces the concept of item into its work flow with the possibility to bind it with any external documentation which can be uploaded into the Medini Analyze work flow.
Within a single project one can manage multiple items, integrating them within the architecture model. Moreover, functions can be added to each item. For example, a cruise control system works by measuring the speed of the vehicle, by estimating the inclination of the driving surface, and by interacting with the vehicle's engine management system. Hence, hazard analysis is made for the chosen single item or for one of its functions. During hazard analysis, the tool guides the user into providing description of the operational situation, the item operation modes, the hazards and the possible malfunctions of the item, so as to obtain a hazard list for the item (or function). For example, for cruise control an hazard could be the unintended acceleration of wheels on a slippery surface without changing the actual speed of the car. Prerequisites, conditions, potential effects of hazardous events can be described.
After the hazard analysis has been completed, the tool provides a wizard for determining ASIL levels. Then, as prescribed by ISO 26262, the tool supports the insertion of a safety goal for each dangerous hazard, along with the inheritance of ASIL levels. A safety goal is a top-level safety requirement, that is defined as a result of hazard analysis and risk assessment and that can be shared by different hazard list entries.).
For example, a safety goal for the cruise control hazard associated with driving on slippery surface can be "do not allow use of cruise control if driving surface is wet". This safety goal inherits the ASIL level of the corresponding hazard.
Afterwards, safety goals can be broken down to functional safety requirements and this is supported with a palette based system to edit SysML structural elements. Figure below shows some possible safety requirements for the cruise control as an example.

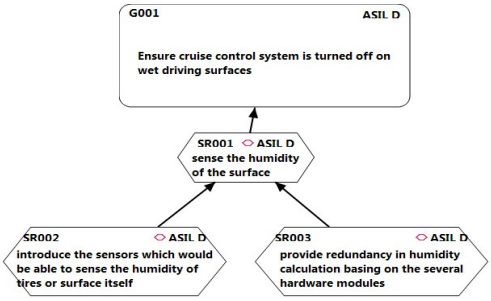


Figure 2. Safety requirements for cruise control by means of Medini Analyze.

The tool embeds support to perform both qualitative and quantitative FTA and FMEA, to facilitate the derivation of requirements. Therefore, it allows the user to create SysML models with the help of an embedded SysML editor. Different levels of abstraction can be created for the defined system, e.g. item level and system level. Architecture SysML models can be used for allocating safety requirements and serves as a base for hazard analysis.
A notable feature of the tool is that it also provides integration with the most popular modeling and development environments in the automotive industry, i.e the Mathworks Matlab/Simulink/Stateflow suite. This is done by linking safety requirements to structural blocks and by the opportunity to see Simulink models inside Medini Analyze.

Traceability
During a system development process it is crucial to have a clear picture of how elements such as requirements, functions, etc are connected to one another and the standard requires these connections to be traced. Because of the complexity of the systems under analysis, shifting the attention from a certain element to a connected one can be not an easy action. Medini Analyze adopts the concept of trace matrix for representing and setting bindings between elements and sub-elements. On the base of this matrix focus is shifted in one-click between connected elements.

Validation
As ISO 26262 sets some rules during activities, there is the need to verify that such rules are fulfilled during the system development process. Let us consider for example *ASIL decomposition*. This is an activity aimed at reducing the likelihood of systematic failures, consisting of substitution of a safety requirement with high ASIL with redundant requirements that have lower ASIL levels. This decomposition can be easily done in a wrong way. To prevent such mistakes a validation engine has been implemented in the tool, so that either the whole project or part of it can undergo the validation to found inconsistencies. The engine makes use of the OCL language, which is a declarative language for describing rules applied to UML models. Therefore, besides the set of rules provided with the tool, users can provide their own rules and/or customize the ones built in the tool.

Document generation.
ISO 26262 requires the production of many documents. Medini Analyze facilitates the production of such documents by generating some of them directly from underlying models and their associated information. For example, "functional safety concept" (the document consisting of safety goals, and safety requirements for these goals) is generated in a completely automatic way by Medini Analyze.

Interfacing and work flow.
Almost all of the file formats used by Medini are XML/XMI, which makes them easy to be included in import/export operations by a wide spectrum of external tools.
For the sake of comfortable user experience, all the activities are made in such a way that they can be used in an iterative manner, without following a predefined sequence of actions and with high degree of independence.
Though Medini Analyze is an Eclipse based platform and extension for other operating systems is planned, at the moment it is available only for the Windows OS.

3.2. EDONA

The EDONA project is a French initiative that aims at constructing an inter-operable integration platform for automotive software development tools allowing the co-development through formalized interfaces over the entire development cycle, rather than providing new tools. The aim is also to integrate safety-based innovations into a common software platform, considering AUTOSAR prescriptions.
The project is directed by Renault, and federates 32 innovating technologies and 13 common source (open-source for members) projects. The project is going to finish in October 2010.
The EDONA components can be generally divided into two parts: Eclipse-based and non Eclipse-based. As a special class of components in EDONA come AUTOSAR components. In Figure 3, the main components (technologies) are shown.

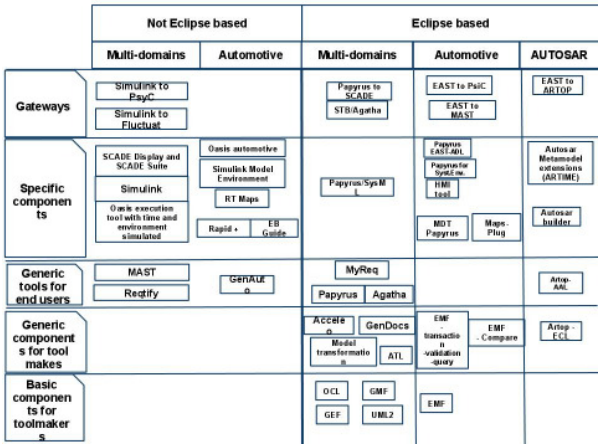


Figure 3. EDONA components.

The framework reuses a set of basic Eclipse components, such as the Eclipse Modeling Framework (EMF) [5] model repository and the ATL Transformation Language (ATL) [6] for model transformation. EMF is a modeling framework and code generation facility for building tools and other applications based on a structured data model, while ATL is a model transformation language and toolkit. In the field of Model-Driven Engineering (MDE), ATL provides ways to produce a set of target models from a set of source models.

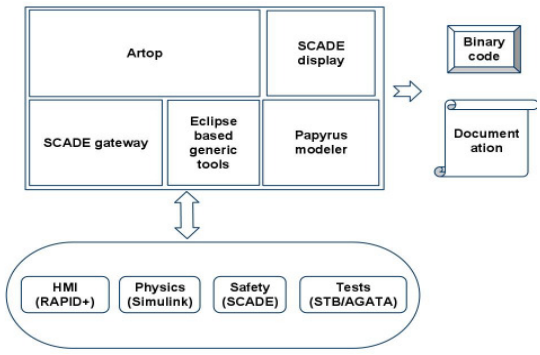


Figure 4. EDONA structure.

The EDONA project aims at integrating tools for the development of embedded software, while addressing dependability, by taking into a particular consideration real-time control systems. This implies requirement and specification traceability, functional and timing validation, deterministic real-time design, as well as HMI (human-machine interface) design and validation with dependability.

For the modelling part, EDONA adopts EAST-ADL and uses Papyrus [13] as the tool for modelling the system.

Safety analysis in EDONA is provided by the Usine Logicielle project. Unfortunately, there is no yet public information about how EDONA covers ISO 26262.

As a means of requirements management, the standalone tool Reqtify [14] and its simplified (as an Eclipse plug-in) version MyReq [14] are used. As an essential activity within the safety critical development process, a large effort has been done towards testing. C code is generated from a Simulink model with the help of the SCADE Suite [8], which is qualified software according to several international safety standards, including ISO/IEC 61508. Then, test cases are generated on the base of source code together with requirements for minimum values of coverage metrics. Test generation is based on Safety Tests Builder IHM [16] and AGATHA [7]. The technique used for test derivation is symbolic automata execution. From Safety Tests Builder IHM, test cases can be passed to a Simulink model for execution or can be exported into an Excel table. For further information about EDONA, see [3].

4. Conclusion

In this paper we have provided an overview of the new upcoming ISO 26262 standard, focusing on the processes and activities that can be supported by tools. Furthermore, we have given a brief overview of Medini Analyze and of the EDONA platform, which are the first initiatives towards integrated software environments for supporting the development of safety-critical electronic components according to the ISO 26262 standard. The aim of these tools is to cover safety-critical design aspects in a richest way in comparison with "one-task" tools, that are already numerous on the market.

Medini and EDONA are just the first steps towards software-based tool support for ISO 26262. It can be expected that more tools will appear on the market in the near future (especially after the official publishing of ISO 26262), and/or that the existing ones will evolve. The target will be to cover the gaps still left in process automation and to get even better integration of the tools composing consecutive tool chains. These targets are being considered by SiSMA, an Italian applied-research project that is just starting.

As it usually happens in software development, a new domain is covered by commercial software firstly, and then, with some delay, open-source solutions appear. Accordingly, new open-source software initiatives giving specific ISO 26262 support can be expected to rise later in the future.

Along with tools, a fast development of safety aspects' support in standardization of automotive system development, which has began to evolve with the publication of AUTOSAR release 4.0, is also expected.

Acknowledgments

We are grateful to the Management Team of ikv+ for a given opportunity to evaluate Medini Analyze with a trial version.

This paper has been written thanks to the support of the SiSMA project.

References

1. ISO 26262 “Functional safety of road vehicles”, 2009.
2. Born M., Favaro J., Kath O. “Application of ISO DIS 26262 in Practice” *Proceedings of the 1st Workshop on Critical Automotive applications: Robustness & Safety*, ACM, 2010, pp.3-6.
3. Ougier F., Terrier F. “Eclipse based architecture of the EDONA platform for automotive system development” *Proceedings of ERTS'2010*.
4. Cuenot P. et al. “Towards Improving Dependability of Automotive Systems by Using the EAST-ADL Architecture Description Language”. *Architecting dependable systems 4. Section: Architectural description language*, 2007, pp.39-65.
5. Budinsky F., Brodsky S., Merks E “Eclipse Modeling Framework” Pearson Education. ISBN:0131425420, 2003
6. Jouault F., Kurtev I “Transforming models with ATL” *Satellite Events at the MoDELS 2005* Conference, Springer Berlin, pp.128-138
7. Bigot C., Faivre A., Gaston C., Simon J. “Automatic Test Generation on a (U)SIM Smart Card” in *Smart Card Research and Advanced Applications*, pp.345-358, 2006
8. Abdulla P., Deneux J.,Stalmarck G., Agren H. and Akerlund O. “Designing Safe, Reliable Systems Using Scade” in *Leveraging Applications of Formal Methods*, Springer Berlin/Heidelberg, pp.115-129, 2006
9. AUTOSAR document. Release 4.0. “UML profile for AUTOSAR”, 2010.
10. AUTOSAR document. Release 4.0. “AUTOSAR Technical Safety Concept”, ID 362, 2010.
11. AUTOSAR document. Release 4.0. Requirements on Methodology, 2010.
12. www.autosar.org
13. www.papyrusuml.org
14. www.geensoft.com/en/article/reqtify/
15. http://sourceforge.net/projects/myreq/
16. http://www.chiastek.com/products/stb.html

Neural Network Algorithms of Gas Turbine Engine Control and Their FPGA Implementation

V.I.Vasilyev
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: vasilyev@ugatu.ac.ru

I.I.Idrisov
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: trikem@gmail.com

A.S.Makarov
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: gttlsh@gmail.com

Abstract[1]

The algorithms of designing the neural network control systems for gas turbine engines (GTE) and their field programmable gate array (FPGA) implementation are considered. The peculiarities of constructing the algorithms of GTE control, checking and fault – tolerance provision with use of neural network are analyzed. The examples of GTE identification and multimode controller design in neural network logic basis and the results of computer simulation are presented. An approach to hardware implementation of two-shaft GTE neural network model on the basis of Altera PLD is discussed.

1. Introduction

One of advanced directions in GTE control is an intelligent control on the basis of artificial neural networks (NN). Various approaches to solving the problems of GTE identification, control and checking with use of NN are considered in [1-4].
The analysis of existing approaches to construction of neural network GTE control systems shows that there are some problems requiring their solution. A list of such problems includes:
- off-line/on-line identification of GTE as multivariable dynamic objects;
- adaptation of GTE control system characteristics to changing engine operation modes, flight conditions, individual peculiarities of particular engine;
- hardware implementation of NN control algorithms on the specialized element base.
In this paper the pecularities of solving the mentioned problems are discussed.

2. GTE identification problem

The purpose of GTE identification is the reproduction of real GTE characteristics as control object by construction of some approximating model connecting the observable data "input/output" and generation on its basis a right strategy of the object control.
At present, an approach based on NN application is widely used for solving this problem [3,4]. A general idea of this approach is based on GTE model representation as a “black box” with NN having the specified structure with a large number of adjusted coefficients (the weights of synaptic connections).
Consider the time series $\{u(t-M),\ldots,u(t-1),u(t);y(t-M),\ldots,y(t-1),y(t)\}$ representing the values of input $u(k)$ and output $y(k)$ vectors for GTE as control object at the particular mode of its operation; $k=t-M,t-M+1,\ldots,t$ - discrete time; $[t-M, t]$ – the observation interval. It is necessary to construct the NN model with the given accuracy reproducing the relationship "input-output" of GTE on the indicated interval of time series.
A general sheme of solving the identification problem takes the following form. The vectors of the object outputs $y(k)$ and NN model outputs $y_{NN}(k)$ are compared under the same input vector $u(k)$. NN training procedure consists in changing its connection weights to reduce the learning error E to the admissible (small enough) value E_{adm}:

[1] **Proceedings of the International Workshop “Innovation Information Technology - Theory and Practice”, September 6th-10th, Dresden, Germany, 2010**

$$E = \sum_{i=1}^{N} \sum_{k=k_1}^{k_1+L} \varepsilon_i^2(k) < E_{adm.}, \qquad (1)$$

where $\boldsymbol{\varepsilon}(k)=\boldsymbol{y}(k) - \boldsymbol{y}_{NN}(k)$ is the residual vector at the *k-th* step; N – the number of the object outputs; $t–M \le k_1 \le t–L$; L – the "time window" width $(L<M)$.

As an example consider the task of designing the NN model of two-shaft GTE. The training sample set consists here of digitized data received by means of on-board recorder, with discretization step 0,5 s. The values of the following engine parametres are used as initial data: G_F - fuel flow rate in combustion chamber (NN model input); n_L - rotation frequency of low pressure compressor rotor (NN model output); n_H - rotation frequency of high pressure compressor rotor (NN model output); T_G^* - stagnation temperature of gas behind the turbine (NN model output). The width of time window forming the training sample set was taken L=80 s, i.e. 160 sequential values by each parameter. For training NN the preliminary processed data were normalized by their reduction to the interval (0,1).

Two NN architecturs were chosen for GTE identification: a)dynamic perceptron; b)Elman recurrent network.

NN training was conducted with use of gradient descent algorithm and Bayes regularization for reducing the probability of overfitting. The idea of regularization principle is based on assumption that the network with smaller values of weight coefficients will possess the best generalization properties. In this case the total learning error can be represented in the following form

$$E = E_1 + E_2, \qquad (2)$$

where E_1 is the error component calculated by formula (1);

$$E_2 = \frac{\lambda}{P} \sum_{i,j} W_{ij}^2, \qquad (3)$$

where λ is the regularizing parameter; $P=M-L+1$ - the size of the training sequence; W_{ij} – the weight coefficient of NN. The more is the value λ, the higher is the degree of regularization (0< λ <1).

For testing the GTE NN model, the input data sequence (observation interval - 2 min 40s, M=320 points, discretization step - 0,5 s) was used.

In Figure 1 the dependences of mean square error (MSE) for trained NN from the number of neurons in the hidden layer are presented (the form of activation function is sigmoid one). As it seen, the best results are achieved for Elman network with 3 neurons in the hidden layer.

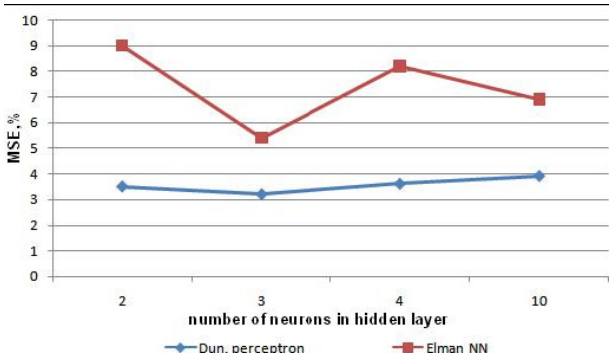
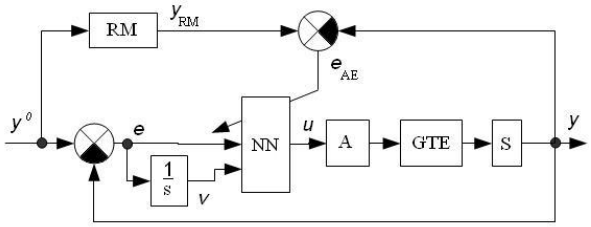


Figure 1. Dependence of MSE from the neurons number in NN hidden layer

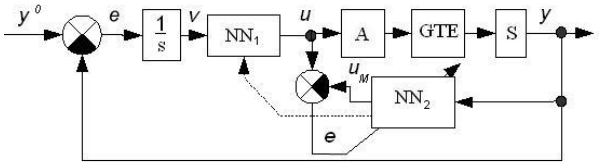
The values of mean square error of NN model at the test sequence by each GTE parameter are: δn_L = 2.19 %, δn_H = 1.36 %, δT_G^* = 2.8 %. More general approach to solving the identification problem on the basis of multilevel regulariztion algorithm is considered in [5]. The obtained results show that NN model adequately reproduces the values of the time series corresponding to real GTE operation processes under the fixed flight conditions.

3. Adaptive multimode NN controller design

Different approaches of NN application in GTE control systems are possible [6,7,8].



a)



b)

Figure 2. Block diagram of GTE control system with NN controller

In Figure 2.a the block diagram of GTE control system with NN controller is presented (A-actuator, S-sensor, RM-reference model). In Figure 2.b the scheme of adaptive GTE control system in a class of NN control systems with identified inverse model is presented. In this case NN (static perceptron with one hidden layer) carries out the function of nonlinear multimode controller providing the generation of control actions to GTE actuator. The input signals of NN are the reference action $y^0(k)$, the control error $e(k)=y^0(k)-y(k)$

and the integral $V(k)$ of control error. The variable u denotes here the output of NN controller by the fuel flow rate in combustion chamber (G_F), y - the measured rotation frequency of low pressure compressor rotor (n_L), y^0 – the reference action by the rotation frequency n_L. The output GTE parameters n_H, T_G* are checked here, but not used directly in the closed control loop. NN controller can be considered as the nonlinear PI controller, the gains of which are adjusted from the condition of the specified control system performance at the fixed set of GTE operation modes.

The necessary condition of GTE control system normal operation is the stability of the closed loop system (Figure 2) at each of the basic modes of GTE operation. It can be shown that the total number of equations (NE) following from the requitement of control system stability provision (the specified replacement of the system transfer function poles at the complex plane) at the given set of GTE basic modes $M_1,\ldots, M_R$ is equal to $(NE) = R{\cdot}N = R(n+1)$, where N is the order of the system characteristic equation; n – the summary order of the object and actuator transfer function; R – the number of GTE basic modes. Assume that there are three basic (steady-state) modes of GTE operation (R=3). Taking into account that the total number of adjusted parameters (NP) of NN controller equals to (NP)=4σ, where σ is the number of neurons in NN hidden layer, the condition of CS correct design (NP) ≥ (NE) can be written in the form

$$\sigma \geq R(n+1)/4. \qquad (4)$$

In our case we have: σ≥3(3+1)/4=3, i.e. the minimum number of the neurons in the hidden layer is equal to three. The total number of the weight coefficients (NP) equals twelve.

As experiments showed, the method of consequent simplex search showed the best efficiency during NN training process. The varied parameters here were the values of NN synaptic connection weights. Taking into account that weights of NN take the values changing in a wide range, preprocessing (scaling) of initial data was necessary for initial simplex construction. As NN learning error, according to requirements imposed for GTE control system, the mean-square criterion of the following form was used:

$$E = \sum_k \Delta n_L^2(k) + 10 \cdot \sum_i \delta T_G^{*2}(k), \qquad (5)$$

where Δn_L is the control error by rotation frequency n_L; δT_G^{*2} - the overshoot value by gas temperature T_G^*.

Tuning of NN weights was carried out with use of the gradient descent method:

$$W_{t+1} = W_t - \alpha \cdot \nabla E(W_t) + m \cdot \Delta W_{t-1} \qquad (6)$$

where ΔW_{t-1} is the weight vector increment at the (t-1)-th step; α – the learning rate; m – the inertia moment.

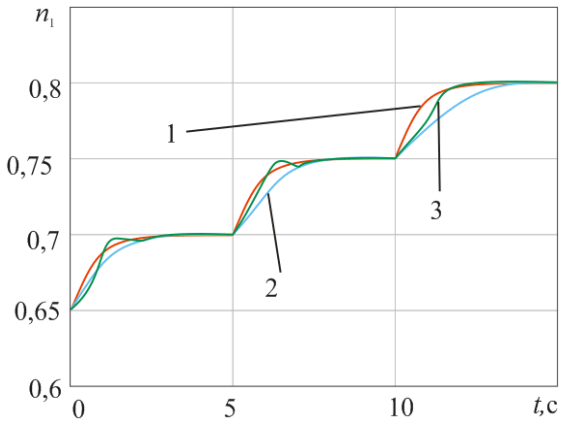


Figure 3. Transient processes by rotation frequency

The transient processes in GTE control system with designed NN controller are shown in Figure 3: curve 1 – the desirable transient processes (exit of reference model); curve 2 – the transient processes by rotation frequency n_L at the initial stage of training the NN controller (initialization stage); curve 3 – the transient processes by rotation frequency n_L for NN controller trained in all range of GTE operation modes, for consequent 5 % increases of the reference action. The analysis of obtained results shows that the settled requirements to GTE control performance are carried out, i.e. application of the offered procedure of designing NN multimode GTE controller is effective.

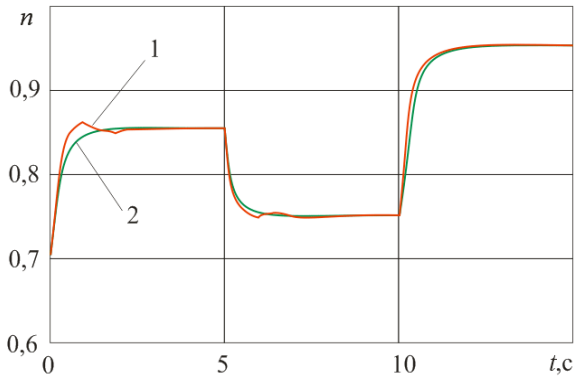


Figure 4. Plot of transients by rotation frequency

In Figure 2.b, investigated system needs the preliminary training of inverse GTE model constructed on the basis of NN2. The parameters of NN1 forming the control actions to GTE actuator, are copied from NN2.

As the analysis shows, the algorithm of NN parameters adaptation is expedient for including during the transients, i.e. at occurrence of a considerable mismatch by rotation frequency n. Each step of adaptation algorithm includes replenishment of data of

training sample, after training NN2, copying parameters in NN1 every second.
In Figure 4 the plots of transients in GTE control system, where curve 1 is the transient by rotation frequency n, are represented; curve 2 – transient on an exit of reference model. From Figure 4 it is visible that after each inclusion of adaptation algorithm the evershoot value by rotation frequency n and transients time decrease.

4. FPGA implementation of NN model of two-shaft GTE

Artificial NN have gained a lot of popularity over the last two decades. These networks can effectively model complex non-linear relationships between inputs and outputs and are also widely used to control complex process. In most cases, an ANN is an adaptive system that changes its network structure and parameters based on external or internal information that flows through the network. The network has to be trained using a training dataset to correctly model the input - output mapping. This training is generally a trial and error based method using an objective function to minimize the error with multiple iterations. With offline training algorithms, an optimal network once found is frozen and deployed in production systems. But dynamics of many real world problems demand an on-line training and adaptation of the network to adjust to the changing input data. The inherent parallelism in these networks is best exploited with custom digital hardware designs which can achieve much higher training as well as recall speeds, compared to the equivalent software implementations. But the lack of a general solution to determine the network structure and parameters and the required regular updates in the trial and error based training process, make these networks difficult to implement in static digital hardware like the ASICs (Application Specific Integrated Circuits). Higher routing requirements for popular fully connected multilayer perceptron models of artificial neural networks incur expensive redesigning, re-routing and fabrication penalties with addition of even a single neuron in one of the hidden layers. To gain the advantage of higher speedups associated with hardware implementations and the flexibility of software implementations, programmable logic devices (PLD)s seem to be a more suitable choice. PLD implementations can offer significant speedups over software implementations with the flexibility of runtime hardware reconfiguration.

Over the last few years, advances in PLDs have resulted in the commercialization of field programmable gate arrays (FPGA) which allow to put large numbers of programmable logic elements on a single chip. The size and speed of those circuits improve at the same rate as 'microprocessors' size and speed, since they rely on the same technology [9-11]. Even with the last decades exponential growth in performance of integrated circuits, many neural network applications still demand increased hardware speed. The main advantage of PLD is their possibility of adjustment on specified functions implementation set by the user. Modern PLD are characterized by low cost, high speed, considerable functionality, reprogramming multiplicity, low power consumption, etc. The supply voltage of PLD with each new generation of these chips permanently decreases that allows us to reduce power consumption at speed increase.

PLD technological features make them convenient object for usage as NN hardware element base and allow us to create chip integrated neural system on one PLD with cascading opportunities. Let us consider the hardware implementation (Figure 5) of NN model of two-shaft GTE.

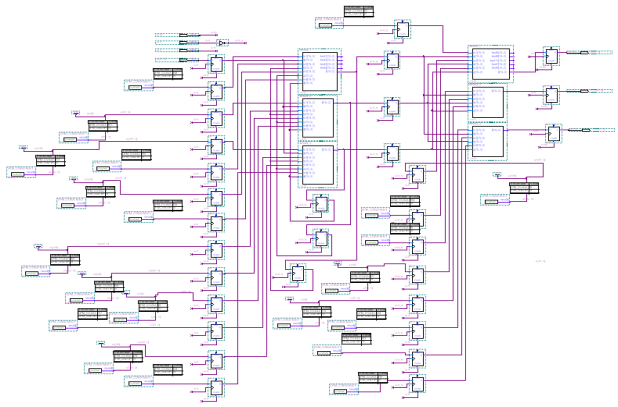


Figure 5. Two-shaft GTE NN model implementation scheme

This NN model belongs to a class of recurrent Elman's networks with time delay units. The network has one input (G_F – fuel flow rate in combustion chamber) and three outputs (n_L – rotation frequency of low pressure compressor rotor; n_H – rotation frequency of high pressure compressor rotor; T_G* - temperature of gas behind the turbine). The neurons of input, hidden and output layers are connected among themselves by synaptic connections with the weights.
Direct implementations of nonlinear activation functions can be expensive in terms of resource costs. A more practical approach would be to use a piecewise-linear approximation of non-linear activation function or lookup table (LUT) approach which essentially pre-stores the output values of the activation function in memory. Advantages of using a LUT based approach is simplicity in design, flexibility to use any activation function depending on the application at hand without hardware redesign.
A floating-point representation of data (weights, biases, inputs and outputs) in a neural network may still be impractical to implement on PLDs despite the current advances in FPGA technology. Thus fixed point implementations are preferred over floating point

implementations as these can use regular integer multipliers and adders to compute the results. The choice of data precision is guided by the implementation cost in terms of area, simplicity of design, speed and power consumption. Higher precision will lead to less quantization error in the final implementation, while lower precision will produce more compact and faster designs consuming less power. A judicious choice needs to be made depending on the application at hand and available FPGA resources. The choice of data precision used would also be driven by the size of the activation function lookup table. Higher precision would require a much bigger table.

It is supposed that all numbers, both input and output for NN, should be in the range [-32768; 32768]. It is meant here that the last two digits concern a fractional part of the number. The input, outputs and the weights of NN are presented by 16 bit code in a binary format. The first bit is assigned for the sign (0 – positive, 1 – negative), other 15 bits are assigned for the absolute number value. Before feeding to the input the number should be multiplied by 100, because it is possible to submit to the input only the integers. However, at the output of multiplication block the number with four digit after comma turns out, therefore it is necessary to fulfill the transformation of the number order obtained at the block output. Order transformation in the multiplication unit is made in order to the addition and activation function calculation occupied less logical cells of PLD. The input neuron has 32 input channels which can be connected to physical pins or an internal interface using the configurable logic array of the host FPGA. The output neuron is similar in both architecture and operation to the input neuron.

For construction of NN model of two-shaft GTE by means of CAD Quartus, the standard blocks (modules) of three- and four- input neurons were designed (Figure 6). Four-input neuron implementation scheme includes four multiplication blocks, four addition blocks, LPM-element for the shift assignment, and also the activated function information block. Let us mark that usage of LPM-macrofunctions allows us to reduce the number of occupied logical units of PLD and therefore to reduce the development cost.

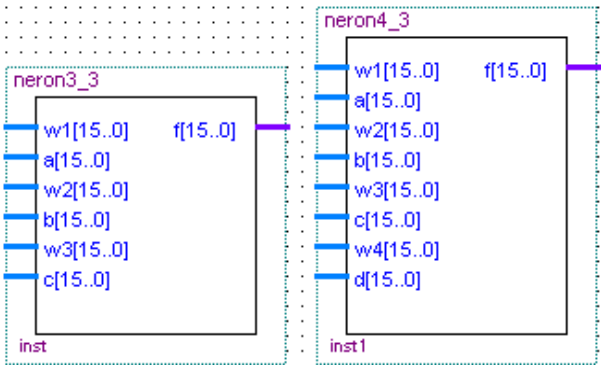


Figure 6. Three and four input neuron representation scheme

Inputs of four-input neuron, e.g., move to the inputs a[15.0], b[15.0], c[15.0], d[15.0]. The synaptic weights of the neuron move to the inputs w1[15..0], w2[15..0], w3[15..0], w4[15..0]. The output signal of the neuron acts from the output f[15..0]. Let us mark that all blocks, designed at the various levels of NN model decomposition, are brought in the standard modules library.

On the basis of developed scheme, simulation of GTE NN model was fulfilled. The input signal test sequence (G_F) throughout 90 steps was used. On the clocking input of PLD the clock rate CLK voltage from the external quartz generator is fed. The clock rate was chosen for the given scheme experimentally, taking into account the requirements to the speed of NN as a whole.

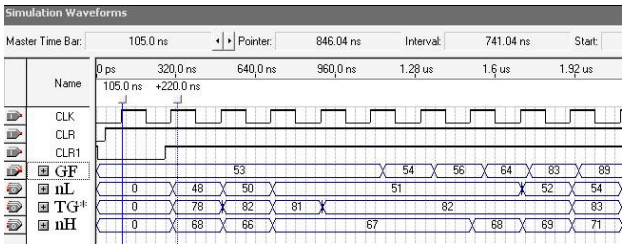


Figure 7. The simulation waveforms of two-shaft GTE in CAD Quartus II

The comparison of two NN operation results was conducted: NN model of GTE realized by software, and the NN model realized by PLD. The mean square error of GTE parameters calculation by means of NN model at the test data series was respectively: δn_L = 1.97%, δn_H = 1.48%, δT_G* = 1.62%. The transient processes in GTE obtained with the aid of NN, realized in CAD Quartus, are presented in Figure 7.

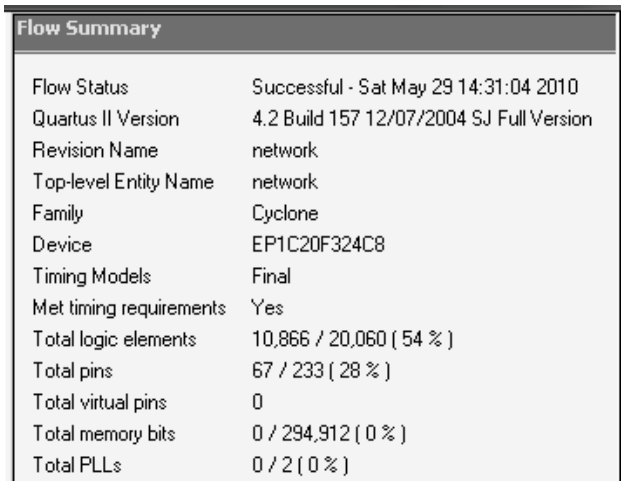


Figure 8. The flow summary of designed NN

Therefore, the hardware implementation of NN models on PLD allowed us to receive the outputs values practically coinciding with the similar values of NN realized by software. Besides, for considered above of

NN model designed in CAD Quartus, the output GTE parameters calculation time for preset value G_F was 220ns. The designed NN occupies 10866 logical elements and 67 pins (outputs) of EP1C20F324C8 (PLD of Altera corporation family) (Figure 8).

Thus, the approach is to develop models of NN-based FPGA can constructively to solve problems related to the implementation of arithmetic operations on real numbers with addition and multiplication units with the order move, with the implementation of the neuron activation function by determining the table function, allowing us to minimize chip resources of the occupied neural network.

5. Conclusion

It is shown that received NN models are suitable for their use as a part of GTE control system model. The dependences of identification accuracy for NN architecture, structure and training algorithms were investigated. The offered technique allows us to solve the identification problem for individual GTE with MSE equal to Δn_L = 2.19%, Δn_H = 1.36%, ΔT_G^* = 2.8%. It is shown that at use of NN designed on the basis of offered technique GTE control processes satisfy the specified requirements in wide range of the engine operation modes. Hardware-realized GTE NN model provides the calculation accuracy of the engine parameters with mean square error less than 2% that meets the specified requirements.

References

1. Toprak, S., Erkmen, A.M., and Akmandor, İ.S.: Identification and Control of a Radial Turbojet with Neural Network and Fuzzy Logic, AIAA'98 36, Int. Aerospaces Meeting and Exhibit, (1998)
2. Zilouchian A., Jamshidi M. Intelligent Control Systems Using Soft Computing Methodologies, CRC, (2001)
3. Dynamic Modelling of Gas Turbines: Identification, Simulation, Condition Monitoring and Optimal Control / Eds: G. Kulikov, H. Thompson, Springer-Verlag, N.Y., (2004)
4. Advanced Multivariable Control Systems of Aeroengines /Eds.: Sun Jianguo, V.I. Vasilyev, B.G. Ilyasov, BUAA Pub. House, Beijing, China, (2005)
5. Vasilyev V., Nugaev I.: Neural Network Smoothing of Geonavigation Data on the Basis of Multilevel Regularization Algorithm, Proc. of the 15th Intern. Conf. on Neural Information Processing, ICONIP'2008, Auckland, New Zealand, 25-28 Nov.2008, pp.109-110 (2008)
6. Sigeru O., Marzuki B. K., Rubiyah Y.: Neuro-Control and its Applications (Advances in Industrial Control), Springer-Verlag, (1996)
7. Werbos P.J.: Neural Networks for Control: Research Opportunities and Recent Developments, http://www.werbos.com/Neural/CDC02.htm
8. Wang T.: A Cooperating Intelligent Control System Based on Expert Systems and Neural Networks, Proc. of the 2nd Intern. Conf. on Intelligent Systems Engineering, Hamburg-Harburg, Germany, 5-8 Sept. 1994, p.171-175, (1994)
9. Reyneri L.M.: Implementation Issues of Neuro-Fuzzy Hardware: Going Towards HW/SW Design, IEEE Trans.on Neural Networks, Vol. 14, No. 1, Jan. 2003, pp. 176-194, (2003)
10. Beuchat J.-L., Muller J.-M.: Modulo M multiplication-addition: Algorithms and FPGA Implementation, ELECTRONICS LETTERS, May 2004, Vol. 40, No. 11, (2004)
11. Altera Corp.: FPGA design security solution using .MAX II devices, (2004)
12. Altera Corp.: Stratix III design handbook, (2006)

Automated Information System of CNC-machines Productivity Loss Analysis

Alexey Anoshin
Branch of Izhevsk State Technical University
Votkinsk, Russia
e-mail: vtkbox1530@udm.net

In most cases, increase in productivity in domestic enterprises is held back by ineffective management and non-optimal operating procedures. Lack of information about machine usage, and equipment status along with inadequate qualification or unorganized staff work makes increasing of production cycle period. Thus, mistakes in organization, maintenance and machine utilization reserve a great deal of idle resources. This leads to a significant rise in production costs, increase in preparation and operating time and, consequently, reduce of the CNC-machines efficiency [1].

Automated Information System (AIS) of performance losses monitoring and analysis is proposed to identify ways of CNC-machines effectiveness increase.

Structurally AIS is a distributed system of data acquisition with centralized storage and processing of information and consists of two parts: data acquisition subsystem, data storage and processing subsystem (Fig. 1).

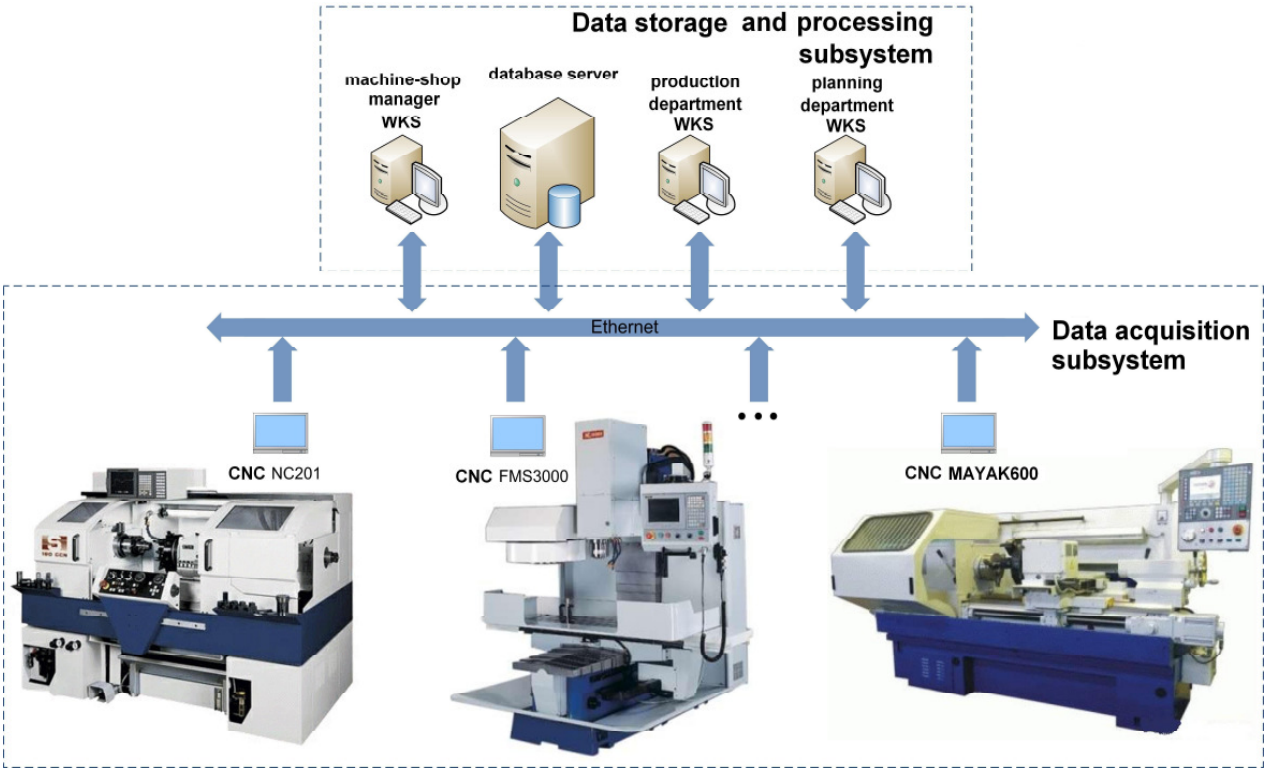


Figure1. AIS block diagram

Data acquisition subsystem functions in the form of a parallel software process on CNC systems of PCNC category [2]. Widely used systems Sinumeric, FANUC, FMS3000, NC, Mayak600 are referred to these CNC-machines. Operation algorithm of this subsystem is as follows.

When running a CNC-machine, data acquisition subsystem records date and time of the machine switching on, makes operator authorization, requests equipment operating conditions and work piece code. Next, continuous control mechanism and recording of machine actuating devices work starts up.

Meanwhile equipment idle time timer is simultaneously initialized. Its reset is carried out by performing of any travel operations on the machine. If the timer is counting time for more than five minutes that exceeds cycle loss time and testifies equipment out-of-cycle idle time, the CNC system is locked and

the operator is offered to indicate the reason of the machine shutdown from the list of out-of-cycle losses given in the table.

Table

Out-of-cycle losses of CNC-machines productivity

№	Denomination of out-of-cycle losses	Designation	Group of losses
1.	Tool setup	$t_{t.1}$	Losses connected with the tool and its adjustment
2.	Absence of tool	$t_{t.2}$	
3.	Toolchange due to breakage	$t_{t.3}$	
4.	Toolchange due to wear	$t_{t.4}$	
5.	Tool verification	$t_{t.5}$	
6.	Waiting of adjuster	$t_{t.6}$	
7.	Repair of machine mechanisms	$t_{r.1}$	Losses connected with machine repair and maintenance
8.	Repair of machine electrics	$t_{r.2}$	
9.	Machine lubrication, coolant refill	$t_{r.3}$	
10.	Waiting for mechanics	$t_{r.4}$	
11.	Waiting for hardware engineer	$t_{r.5}$	
12.	Waiting for electricians	$t_{r.6}$	
13.	Cleaning at the end of shift	$t_{o.1}$	Organizational losses
14.	Periodical cleaning	$t_{o.2}$	
15.	Delivery of finished parts	$t_{o.3}$	
16.	Getting of work pieces	$t_{o.4}$	
17.	Work negotiations	$t_{o.5}$	
18.	Turning over the shift	$t_{o.6}$	
19.	Lack of work pieces	$t_{o.7}$	
20.	Personal necessities	$t_{o.8}$	
21.	Absence of job assignment	$t_{o.9}$	
22.	Absence of program	$t_{o.10}$	
23	Rejects control at the adjustment	$t_{c.1}$	Losses connected with control of parts and work pieces
24	Rejects control at malfunction	$t_{c.2}$	
25	Substandard materials control	$t_{c.3}$	
26	Finished parts control	$t_{c.4}$	
27	Switch-over to new product	$t_{s.1}$	Losses due to the machine switch-over
28	Switch-over to new operating procedure	$t_{s.2}$	
29	Switch-over to new tool	$t_{s.3}$	
30	Switch-over to new equipment	$t_{s.4}$	

Thereby timing of machine and operator work is realized, as well as recording and identification of equipment idle time causes.

Then collected information over Ethernet network is transmitted to data storage and processing subsystem, where indices of cycle, technological and actual performance are calculated. Coefficients of equipment utilization and availability are calculated.

Thus, the cyclic performance indicator allows assessing of the machine effectiveness in comparison with similar equipment operation at processing of one work pieces batch [3]. In this case assessment of performance is done by number of manufactured useful products for the time period T and is calculated by the following relationship:

$$Q_p = \frac{1}{t_w + t_i} = \frac{p}{T}$$

where t_w - time spent on working travels; t_i - time spent on idle passes, spindle-on/off, tool change, preparation of parts, etc., T - period of time to process p-products.

Application of this indicator is suitable for series production. For small-scale and limited production it is preferable to apply a machine use factor as a performance criterion, it determines what part of the planned period of time θ equipment actually operates and puts out products

$$\eta_u = \frac{\theta_w}{\theta_w + \sum \theta_i}$$

where θ_w - total time of equipment work during the period of time θ; $\Sigma\theta_i$ - total idle time of equipment during the same period of time.

Thus, index of actual performance is determined by relationship:

$$Q = Q_p \eta_u$$

For heads of an engineering factory in most cases it is not enough to know the actual performance of the equipment and the indices of its use for effective management. It is necessary to decode, why a machine use factor η_u takes one or another value, what causes and types of idle times are prevalent. In this case partial factors of η_u, reflecting the impact of certain idle time types, may be informative:

$$\eta_u = \eta_{tech}\eta_{sw}\eta_{cap},$$

where - η_{tech} technological use factor, numerically representing fraction of time during which the equipment, provided with everything required, works; η_{sw} - switch-over factor shows fraction of planned time period, during which provided with everything required the equipment can work, releasing a certain product; η_{cap} - capacity factor shows what fraction of planned time period the equipment is provided with everything required to work, i.e. how it is employed in the particular conditions of production.

Calculation of these factors data is carried out by relationships:

$$\eta_{tech} = \frac{\theta_w}{\theta_w + \sum\theta_t};$$

$$\eta_{sw} = \frac{1}{1 + \frac{\sum\theta_{sw}}{\theta_w + \sum\theta_t}};$$

$$\eta_{cap} = 1 - \frac{\sum\theta_{org}}{\theta};$$

$$\sum\theta_t = \sum\theta_{tool} + \sum\theta_{rep.} + \sum\theta_{cont.};$$

$$\sum\theta_{tool} = \sum_{i=1}^{n} t_{t.i};$$

$$\sum\theta_{rep.} = \sum_{i=1}^{m} t_{r.i};$$

$$\sum\theta_{cont.} = \sum_{i}^{k} t_{c.i};$$

$$\sum\theta_{org} = \sum_{i=1}^{l} t_{o.i};$$

$$\sum\theta_{sw} = \sum_{i=1}^{p} t_{s.i}.$$

where $\Sigma\theta_t$ - own or technical idle time (time spent on replacing and corrective adjustment of tool, repair and maintenance of machines, etc.); $\Sigma\theta_{org}$ - organizational idle time caused by external factors not related to technology and machine design; $\Sigma\theta_{sw}$ – idle time associated with equipment switch-over; $t_{t.i}$, $t_{r.i}$, $t_{o.i}$, $t_{c.i}$, $t_{s.i}$ - out-of-cycle equipment losses of the corresponding group (see Table.); n, m, k, l, p – number of idle time types in the corresponding group.

The advantage of actual performance assessing with relative factors η_u, η_{tech}, η_{sw}, η_{cap} is clearness and simplicity of numerical values interpretation. For example, if a CNC-machine is used according to the production program by 90% (η_u = 0.9), in addition 10% of the time left is idle for switch-over (η_{sw} = 0.9), and in the periods of time provided with everything required works only 80% of the time (η_{tech} = 0.8), then as a result fraction of planned time period when the CNC-machine works and puts out products, is

$$\eta_u = \eta_{tech}\eta_{cap}\eta_{sw} = 0.8 \cdot 0.9 \cdot 0.9 = 0.65$$

i.e. the potentialities of the CNC-machine are used by 65%.

In addition, partial factors can also be represented as the product of idle-time coefficients of more detailed classification:

$$\eta_{tech} = \prod_{i=1}^{n}\eta_{u.i} \cdot \prod_{i=1}^{m}\eta_{r.i} \cdot \prod_{i=1}^{k}\eta_{c.i},$$

$$\eta_{cap} = \prod_{i=1}^{l}\eta_{o.i},\ \eta_{sw} = \prod_{i=1}^{p}\eta_{s.i}$$

where

$$\eta_{u.i} = \frac{\theta_w + \sum_{j=0}^{i-1} t_{t.j}}{\theta_w + \sum_{j=1}^{i} t_{t.j} + t_{t.i}},$$

$$\eta_{r.i} = \frac{\theta_w + \sum\theta_{tool} + \sum_{j=0}^{i-1} t_{r.j}}{\theta_w + \sum\theta_{tool} + \sum_{j=1}^{i} t_{r.j} + t_{s.i}},$$

$$\eta_{c.i} = \frac{\theta_w + \sum\theta_{tool} + \sum\theta_{rep.} + \sum_{j=0}^{i-1} t_{c.j}}{\theta_w + \sum\theta_{tool} + \sum\theta_{rep.} + \sum_{j=1}^{i} t_{c.j} + t_{c.i}},$$

$$\eta_{s.i} = \frac{\theta_w + \sum\theta_t + \sum_{j=0}^{i-1} t_{s.j}}{\theta_w + \sum\theta_t + \sum_{j=1}^{i} t_{s.j} + t_{s.i}},$$

$$\eta_{o.i} = \frac{\theta_w + \sum\theta_t + \sum\theta_{sw} + \sum_{j=0}^{i-1} t_{o.j}}{\theta_w + \sum\theta_t + \sum\theta_{sw} + \sum_{j=1}^{i} t_{o.j} + t_{o.i}}.$$

That is, a group of coefficients $\eta_{u.i}$ shows a fraction of time during which the equipment provided with everything required works, and what fraction of time the equipment is idle because of change, control, or absence of tools. And in its turn, set of coefficients $\eta_{r.i}$ allows assessing fraction of equipment idle-time period with faultless and adjusted tools during repair of particular mechanisms or machine components.

Group $\eta_{c.i}$ assesses fraction of idle-time period of faultless equipment with adjusted tools due to control of parts and pieces. In this case $\eta_{u.i}$, $\eta_{r.i}$, $\eta_{c.i}$ do not take into account equipment switch-over time.

Group $\eta_{s.i}$ allows defining a fraction of switch-over time in periods when the equipment has all organizational preconditions for work, i.e. there are work pieces, tools, electricity, etc.

And finally, group $\eta_{o.i}$ shows what fraction of planned time period equipment is provided with everything required for work, and what fraction of time period equipment is idle due to relevant organizational reason i^{th}.

Values of these factors and parameters can be represented both in integral, and in expanded form. In addition, parameters, presented in expanded form, make it possible to monitor trend in productivity losses, both for individual machines and for shop or the whole production.

Integral values of the parameters represented in the form of pie charts allow estimating graphically the fraction of equipment idle-time by particular reasons.

Using this system of factors and parameters jointly with developed software, we carried out performance losses analysis of five CNC-machines at one of the machine-building enterprises of town Votkinsk. Experimental operation results are represented in Figure 2.

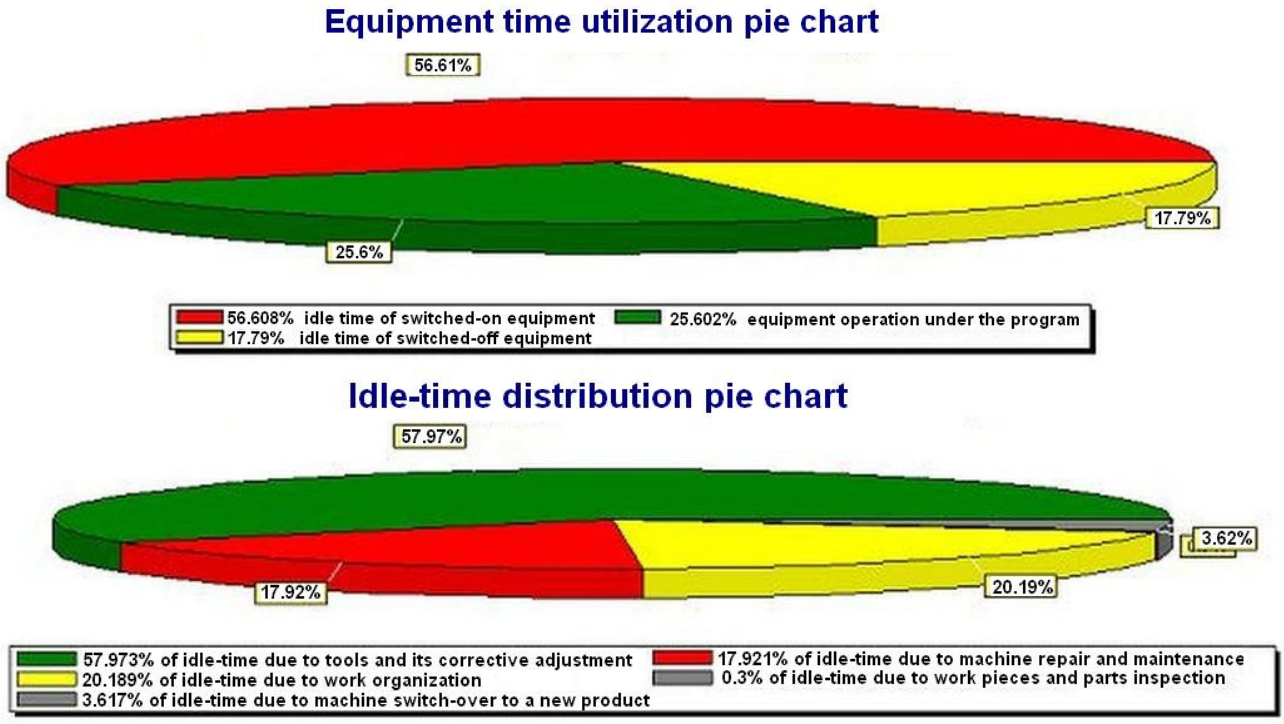


Figure 2. The results of AIS experimental operation

Analysis of equipment time utilization pie chart shows that only 25.6% of working time the equipment operated under the program and put out products. Moreover cycle and out-of-cycle idle time of switched-on equipment accounted for 56.6%, and out-of-cycle idle time of switched-off equipment were 17.8%. Out-of-cycle idle time of switched-off equipment is related to either repair or absence of operators that is controlled by the enterprise scheduled service.

Idle-time distribution pie chart of switched-on equipment makes it possible to examine out-of-cycle productivity losses. Thus, 57 % of idle-time is due to either absence or corrective adjustment of tools, 18% of idle-time is due to machine repair and maintenance, 20% of idle-time is organizational losses.

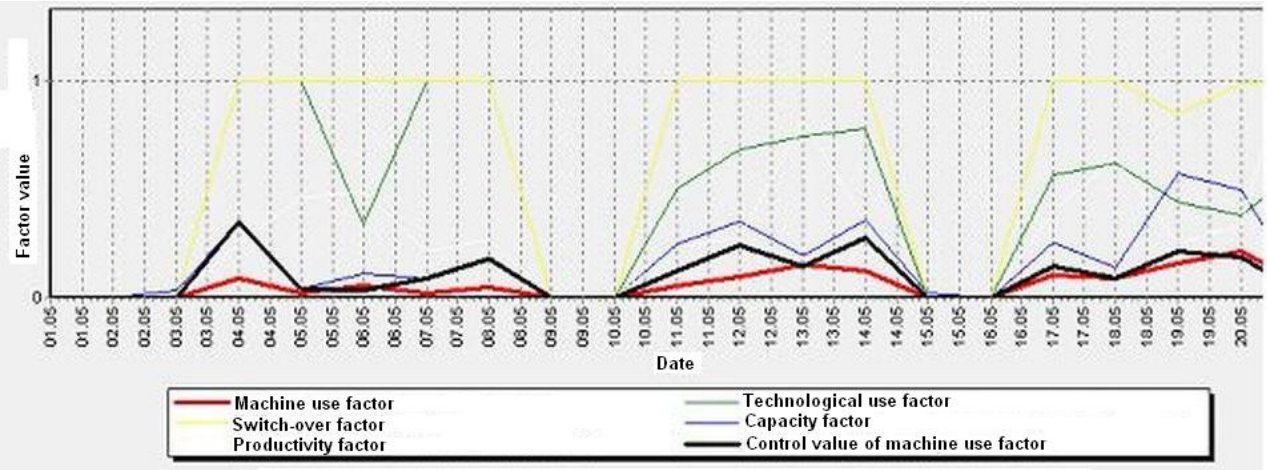


Figure 3. Bar-and-pic chart of factors

Values of factors characterizing all types of losses and their time distribution are shown in Fig. 3.

Here it is clear that value of equipment use factor does not exceed 0.3 – that indicates low efficiency of equipment use, and inadequate preproduction of this enterprise.

Thus, experimental operation of automated information system made it possible to evaluate specific production effectiveness and perform analysis of reasons of equipment out-of-cycle idle-time. Thereby, main out-of-cycle idle-time of switch-on equipment for this enterprise is due to low-quality tools or mismatching of cutting conditions. Organizational idle-time and idle-time due to repair of badly worn equipment also make a significant part. Readjustment and part inspection losses do not exceed 4%.

With these data head of the enterprise has the opportunity to choose the most effective ways of CNC-machines productivity improvement, using available manufacturing resources more efficiently.

References:

1. Molchanov G.N. "Increasing in efficiency of processing on CNC machines". - M.: Mashinostroenie, 1979. – 204 p, ill.

2. «FMS3000». NC device. Operating manual. Description of electrical automation language. Description of parameters system. Edit. 5.0 – Nizhny Novgorod: «Modmash-Soft» Ltd., 2005. – 116 p.

3. "Automation of production processes in engineering": manual: Two Vol./ Strihladze A.G., Voronov V.N., Boriskin V.P.. – Stary Oskol: TNT, 2008. - Vol. 1. – 148 p.

Towards a more cognitive and cooperative ICT-Infrastructure for Production Networks

Dipl. Wi.-Ing. Ramez Awad
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: ramez.awad@kit.edu

M.Sc.. Stilian Stanev
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: stilian.stanev@kit.edu

Dipl. Inform. Vitalis Bittel
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: vitalis.bittel@kit.edu

Dipl. Inform. Stefan Pölz
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: stefan.poelz@kit.edu

Dipl. Wi.-Ing. Alexander Burger
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: alexander.burger@kit.edu

Dr. Denis Popov
Department of Computer Science and Robotics
Ufa State Aviation Technical University
Ufa, Russia
e-mail: popov@ugatu.ac.ru

1. Abstract[1]

Today's ICT systems can be considered to consist of a large number of various components with different functionalities, which have to collaborate in a meaningful way to provide enhanced functionalities. Managing such a collection of arbitrary, networked components has proven to be a very challenging task especially when considering ad-hoc interoperability, robustness, incomplete, partial or inaccurate information. Therefore future ICT systems components must therefore be aware of their role and function in the system, as to interactively and selectively collect information about the network's context, as well as the limited available resources in the network, and act upon this information. The nodes need to have a concept of self and the knowledge to "understand" this concept – self-awareness. This paper intends to show a model-based three level approach towards describing component's functionalities, capabilities and limitations as well as their variation considering the internal state of the component.

Keywords—Functional Description, Machine Learning, Adaptive Information Fusion, Cognitive Processing.

1. Introduction

Over the last two decades, information and communication technology (ICT) has become an integral part of our life. It has provided us with fast, cheap, reliable communication means for sharing information; it has introduced new society-shaping services like internet banking, and credit cards. ICT played a major role in the research advancements made over the last 10-15 years in genetics, physics, astronomy, nanotechnology, etc. The impact of ICT is much more and further reaching than any other technology invented so far. ICT is empowering society in both the economic and social context [1].

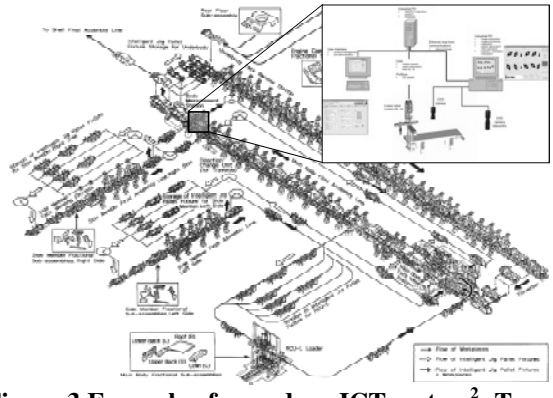


Figure 3 Example of a modern ICT system[2]: Toyota Assembly line [3]

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

[2] In the following context the terms "system" and "network" will be used as synonyms. Consequently, the

However, modern ICT systems are increasingly heterogeneous, ubiquitous and mobile. By necessity they are growing in size, complexity and importance. They consist of a large number of various nodes (embedded devices, software programs, services, platforms, etc.) with different functions. These networked and distributed systems do not perform their tasks in isolation but have to work together in a meaningful way. They have to be able to achieve an efficient use of resources and to cope with the emergence of process mobility and with heterogeneous devices in dynamic contexts, while operating in an ever-changing, dynamic, complex environment, not fully specified a priori [2].

Managing this level of complexity proves to be a very challenging task especially when considering the following [4], [5], [6], [7] and [8]:

- These systems need to provide an infrastructure able to support the coordination and organization of heterogeneous components originating from different, possibly non-trusted, fault-prone, malicious, or selfish sources.
- The nodes in the system have to cope with incomplete, partial and/or inaccurate information.
- The ICT systems need to be robust and able to adapt to environmental changes, to the movement of people and hardware and of data and applications.
- Ad-hoc interoperability with new emerging devices.
- Scalability, large-scale and efficiency issues.

This level of complexity requires a radical shift in perspectives. New approaches for massively concurrent systems are needed. This ranges from new design paradigms, parallel and distributed programming approaches, to artificial intelligence concepts, as manual management of these systems is no longer applicable, as demonstrated by the following examples [9] and [10]:

- A Cisco study found a 5% error rate in manually performed network changes. In a system with 1000 devices, this would mean that 50 devices will be improperly configured at some point. Because of the precise nature of network device configuration, a single error could result in partial or total network failure.
- Another study found that on average the manual recovery time for a network device is 30 min, which in some business translates into literally millions of dollars.
- When a new automotive product is introduced, the IT department of the automotive OEM often requires more than a year to prepare the corresponding ICT systems, in terms of data only!
- Consider a network with 1000 devices (e.g. a small production system) where configuring each machine requires 10 minutes. An administrator would have to work around the clock for a complete week to configure all the devices manually. A week where the entire production system would be out of service.

Future ICT systems and networks are therefore envisioned to operate in an autonomous manner. They may have to replace services that are no longer available, due to certain nodes leaving, with others that offer the same or similar functionalities, and take advantages of new services provided by the network environment, due to new nodes joining the network. They must be "*flexible so as to provide seamless access to located services, extensible so they can be easily modified or upgrade themselves; and interactive so as to cope with highly dynamic environment*" [2].

Systems should in particular have the following features:

a) *Global awareness* regarding performance, network and resource parameters. This global awareness should be used to optimize overall performance and resource usage in response to change, adapting to both context (such as user behaviour) and internal changes (such as the integration of new resources).
b) *Bidirectional influence* between the global and local levels. The global knowledge distributed over all the nodes in the network should be used to influence the local knowledge of the nodes, whereas the localized knowledge of the nodes should be used to create and influence the global knowledge of the network.
c) *Dynamic management* of virtualized and physical resources. Resources in the system should be allocated as needed, by coordination mechanisms allowing the nodes to ask for resources and/or offer own resources for use while achieving a global optimal use of these resources.

Nodes in such future ICT systems or networks must therefore be aware of their role and function in the network, so that they can interactively and selectively collect information about the network's context, as well as the limited available resources in the network, and are able to act on this information.

However, this raises the question: Which information is relevant? What amount of information is necessary and sufficient for the node to act? How does a node or network realize the need for change?

To answer these questions the nodes need to have a definition or concept of "self". This concept of self would then enable the network to define things and experiences as associated with self (experiences occurring within the nodes boundaries) or associated with not-self (perceptions of objects in the environment) [11]. The nodes need to be self-aware, they need to know themselves.

This raises the question: How does a device or node know about itself?

Prof. Hayes provides a very simple answer to that question:

"If something which can represent things needs to know about itself, just give it a way to represent itself

terms "node" and "device" will also be used as synonyms.

to itself." (Pat Hayes, IHMC, U. West Florida, In: CAP-2000)

2. General Approach

While this presents an elegant solution to the problem, the next questions presents itself as to what should be represented. The physical layout, the goals, the "beliefs"...?

There are already systems which are aware of their internal state, e.g. notebooks continuously monitor their battery levels. This is not sufficient for self-awareness. On the other hand, awareness of the "beliefs" of a system would go far beyond the level of awareness needed for autonomous self-aware systems.

Three levels of awareness can be identified as needed for autonomous self-aware systems. A device in the system needs:

1. *Function-based Self-Awareness:* Awareness of its basic functions, capabilities and limitations, as well as the internal state relating to the performance of these functions (Figure 4).
2. *Self-based World-Awareness:* Awareness of the world and environment, as it relates to self, for context and situation awareness (**Figure 5**).
3. *Network-oriented Impact-Awareness:* Awareness of the network for global impact awareness.

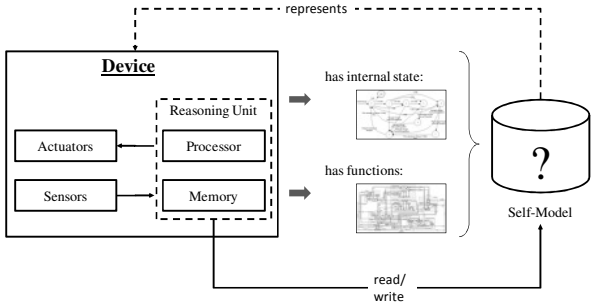


Figure 4 Function-based Self-Awareness

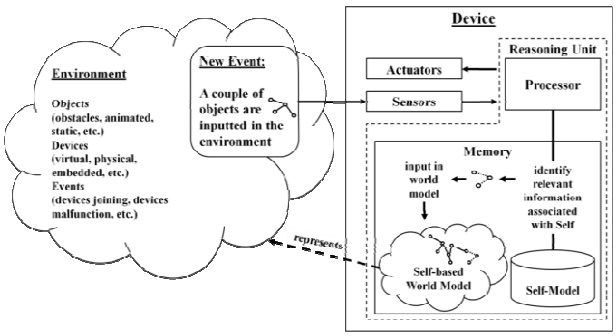


Figure 5 Self-Based World-Awareness

In the following the single levels are elaborated:

- Function-based Self-Awareness: a device needs to be aware of two related information sets. First, its functions and capabilities, so that the device can understand its function and role in the system. A description of their capabilities and limitations provides the devices with the means to react to changes in their environment in line with their role in the system. Second, a device needs to be aware of its internal state as it might affect the performance of its functions
- Self-based World-Awareness: The second level of awareness considers the environmental changes mentioned in the previous level. In order for a device to react to those changes, it needs to be aware of these changes and understand their relevance towards itself. Thus, the devices must have access to sensors or other information sources. Each sensor provides the device with different information. These information need to be fused together in an unified world model.
- Till now only passive aspects of awareness have been regarded. However, to exploit the full potential of self-aware devices, the network-oriented impact awareness has to be included. Impact awareness describes the devices ability to plan different action alternatives, to assess the impact of their actions based on their world model and to select the best alternative to be executed. In a self-aware system (SAS), where the devices must cooperate to execute distributed tasks this entails merging the different subjective world views of the cooperating devices into a more global view of the environment and network. Thus the impact awareness needs to be broadened to a more network-oriented approach, meaning that the global impact of the devices actions on the entire network need to be considered. To that extent, the devices need to create a joint world model by merging their different views of the network in an overly.

The following outlines the model-based approach to tackle the different awareness levels and their interdependencies:

To capture the functions, capabilities and limitations of a device, an ontology-based functional model must be developed, which focuses on the low-level sub-functions and material, information and energy flows within the device [12]. To allow for a certain degree of adaptation, the model will support the aggregation of the sub-functions in a hierarchy (as depicted in Figure 6) to higher level distributed services, even services not perceived during the original design of the global system. The functional model will also include the different parameters that affect the performance of the functions. Using a request-based or broadcast strategy the functional description of a device is communicated to all other devices in the network. Each device will have an integrated multi-level world model of its environment, with different abstraction and detail levels. Large-scale ICT networks or systems result in large quantities of heterogeneous information, with varying accuracy and credibility, which can be obtained through various sources, like sensory systems, databases, GSM networks, world wide web, etc. [13].

In order for the different devices to build an integrated image of their virtual and physical environment in their world models, they need to fuse those information together into a holistic world model. The decision as to which information to fuse will be decided by a reasoning unit, utilizing the functional model. The functional description of other devices will be part of the information collected and fused in the world model. As it is not feasible for each device to have a global view of the entire system, the concept of *neighbourhoods* (physical, task-defined or by similarity) is introduced in the world model, meaning the nodes will have an intense knowledge of those parts of the whole system that are relevant for them (aka neighbourhood) and a coarse knowledge about nodes that are less relevant.

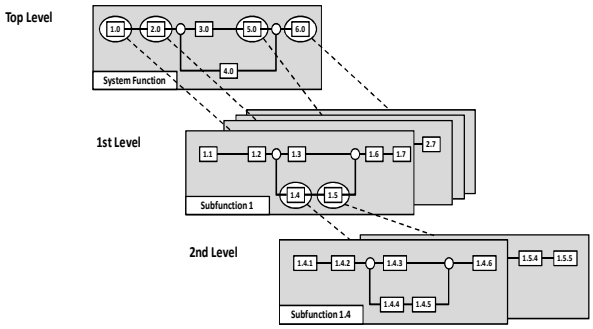


Figure 6 Hierarchical Functional Structure

The generation of world models suitable for the role of the device in the system has be investigated. To that end, a generic world meta-model needs to be specified and methods and tools for the dynamic generation and adaptation of the world model using the functional model and the available information sources need to be developed. This would enable a dynamic model which evolves with every new information source joining the network. For the fusion process itself, an ontology-based approach as outlined in [14] should prove sufficient.

The impact-awareness is achieved via the reasoning unit. When a task is announced in the system, the reasoning units of the devices decompose the task into subtask, and then map those subtasks to the different functions in its functional model, thus generating an action plan. Then the reasoning unit uses the world model to estimate the impact of its actions and selects the action plan with the most desirable impact. After executing the action plan the reasoning unit identifies the effects of its actions as changes to the world model. By comparing these changes to the desired impact the reasoning unit can identify potential differences and deduct the possible causes for the divergence. It then can use that acquired knowledge to further improve the impact estimation process for the next task (i.e., it can learn from experience).

Since the devices cooperate to execute distributed functions we must broaden the cognitive process described in the previous paragraph to a more network-oriented approach. In this scenario the reasoning units of the devices have to communicate as to the distribution of the subtasks to the functions of the different devices. For that they first need to have knowledge of the functions available in the network. As every device has just a limited view of the network, a more global view of the network must be created by integrating the local world models of the devices into a shared world model. To that end, the use of fractals[3] and appropriate aggregation processes shows some potential. Once the subtasks have been distributed the shared world model can be used for a global impact estimation of the generated action plan alternatives. The knowledge acquired by evaluating the actual impact of the action plan against the estimated impact is then used by the reasoning unit in the next iteration, thus closing the loop to a distributed cognitive process.

To implement this distributed cognitive process, a function-based service composition algorithm has to be developed, able to decompose incoming tasks into subtasks and match those subtasks to the different devices in the network given a set of application requirements, available monitoring data, functional descriptions of physical nodes and the current network-context as provided by the world model, while adhering to both local and global system policies. This algorithm is envisioned to utilize the hierarchical aspect of the functional model to develop a virtual overlay network above the physical layer of the devices by aggregating the functions offered by the different devices to a functional unit. To that end, the suitability of different machine learning algorithms for the implementation of the distributed cognitive process can be researched.

3. Dynamic System Considerations

Self-awareness is not suitable or adequate for every ICT system. Self-awareness for a system, that just computes answers, might be feasible but not prudent. Self-awareness makes only sense for systems that have a persistent activity and interact with a highly dynamic environment, e.g. production systems.

This paper identifies the following properties of a Self-Aware System (SAS):

- The SAS is dynamic. Devices are continuously joining and leaving the system.
- The SAS is heterogeneous. It incorporates virtual and physical devices from various domain contexts.
- The SAS behaviour is non-deterministic, as there are too many variables (policy of device, behaviour of other devices, environment, etc.) influencing the decision-making process of the devices.
- In a SAS, devices are aware of their own capabilities and those of other devices around them which may

[3] Model that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole.

provide the networking and system services or resources that they need.

- In SAS, devices can self-organize to execute certain tasks without human interference.
- In SAS, devices have the ability to learn new skills and acquire knowledge.
- In SAS, devices only have partial knowledge of the system.

So when considering the dynamic system (as depicted in **Figure 7**), we are faced with additional challenges, which need to be also addressed:

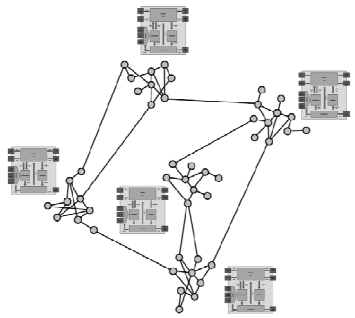


Figure 7 Autonomous Self-Aware System/Network

(1) Because of the complexity and dynamics of modern ICT systems with devices leaving and new devices joining the network, the different relevant information sources are not known a priori and will vary over time, meaning relevant information sources need to be discovered at runtime [13]. Furthermore, these information may correspond to different levels in the world model and may be conflicting. Therefore, the information fusion process will must be adaptive (to accommodate the time-varying availability of heterogeneous information sources), as well as able to manage the information consistency across the different levels. Thus, methods for identifying new information sources using their functional description and incorporating these new information sources in the fusion process at runtime are needed. As the performance of information fusion in ICT systems presents a problem [15], which could affect the responsiveness needed for self-awareness, *SAREA* will pursue a distributed information fusion approach [16] which has greater scalability.

(2) In self-aware systems all entities participate in network control through individual interactions. Communication methods for ad-hoc autonomic networking have been sufficiently researched in FP6 ICT Calls 4 and 5. However, the establishment of trust and security in the context of autonomic networks remains an open issue. Traditional third party or centralized, hierarchical trust models fail because of the non-deterministic behaviour exhibited by autonomic nodes [5]. Furthermore, trust in self-aware systems depends heavily on the role of the entity in the application and network context [17]. This paper proposes the assignment of trust values to devices dependent of their current role in the application and network, as well as the impact of reputation (trust values of nodes as part of collaboration and not as individuals) on the trust between nodes. The reasoning unit will be able to use dynamic role-based trust methods, for distributed and modular trust establishment with partial and incomplete information. The information needed for these methods will be part of the world model and holders of this information will be considered when constructing the corresponding meta-model.

(3) Self-aware networks are part of the future of cognitive networks [18]. By applying the definition for cognitive networks provided by [19]: as networks with a cognitive process that can perceive current network conditions, sense the environment, plan actions according to input from sensors and network policies, decide which scenario fits best its purpose using a reasoning engine, and finally act on the chosen scenario and learn from the consequences of its actions, it becomes evident that knowledge management and machine learning is also an integral part of the self-aware network. In the context of the FET project "NEW TIES"[4] the learning process was divided in individual learning, evolutionary learning and social learning. In networks with high fluctuations concerning the devices and knowledge loss because of devices leaving the network, the importance of social learning rises to a whole new level. At the same time, because of the autonomy and heterogeneity of the nodes in the self-aware network, as well as the incomplete and partial information available to the nodes, traditional centralized machine learning approaches utilizing the hypothesis of complete information are not applicable [20]. A distributed approach is needed. To that goal, the reasoning units need to be embedded with distributed machine learning algorithms. Furthermore, a subroutine for social learning enabling the exchange of explicit knowledge will be integrated in the cognitive process of the reasoning unit.

(4) There are two collaboration activities that could occur in a self-aware network: application collaboration and network collaboration. The network collaboration is responsible for the establishment of the network and the basic communication, transport and organizational layer above the physical layer formed by the devices. The application collaboration on the other hand is responsible for the negotiation between the different devices and cooperation of those devices to perform a distributed task in the context of an application. Traditionally, the two forms of collaboration are separated and allocated to different layers in the OSI model, however the two are so interconnected that

[4] http://www.cs.vu.nl/~gusz/newties/newties.html

this separation is no longer applicable [21], as the internal status of the network could affect the quality of an application and the execution of an application can change the network status. To overcome this disjunction this paper proposes the extension of the information sensing and sharing framework (ISS) developed by the InfoLab21 at the University of Lancaster to a generic framework for cross-layered communication. This new framework "breaks down layer boundaries, share information about network and application state, which allows performance to be optimized" [21]. As to structure the communication within such a cross layered frame work a Formal Notational Language will be specified.

There is a lot of information exchanged between the devices (functional descriptions, sensor information, knowledge, etc.) in a self-aware network, yet there are no mechanisms to ensure that the information exchanged is understood in the same way by all devices. A traditional collaboration ontology approach would only provide a common understanding of concepts defined at the beginning of the network set-up. However in a dynamic self-aware network, new devices utilizing new unknown concepts are continuously joining the network. To meet the challenge of a "common understanding of information" the functional description will be mapped or even directly integrated into collaboration ontology common to all devices. Because of the heterogeneity and dynamic of modern ICT systems and environments, this collaboration ontology must be constructed from the start to be able to automatically incorporate new concepts provided by the functional description of new unknown devices. The meaning of new unknown concepts eludes the collaboration ontology at the start. Therefore, the collaboration ontology, needs to be able to describe the meaning of new concepts as it emerges out of the application context of the new device, and integrate these concepts in the collaboration ontology. Essentially, this touches the field of ontology "merging".

4. Applicable Business Scenario

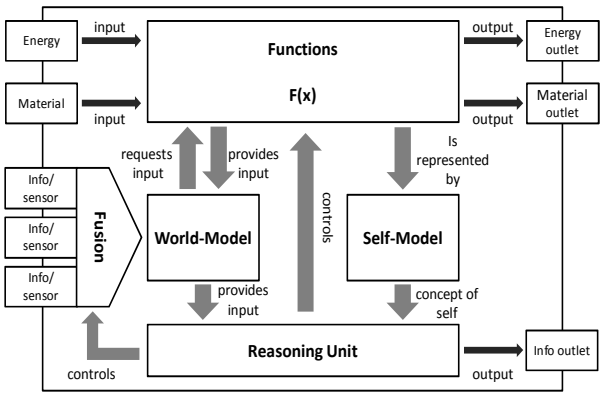


Figure 8 Proposed architecture of a self-aware device/node

As a radical shift in design paradigms (Figure 8) is being proposed here; the proposed research cannot fall back on current devices for testing the concepts, methods and algorithms developed. Therefore, the autonomous self-aware system should be implemented in a Multi-Agent-Platform (MAP).

The development of the MAP will include the conceptualization and implementation of several aspects of a multi-agent platform, such as a plug-and-play functionality for agents, ad-hoc interoperability, wrapper-service for encapsulating new devices joining the network, a service-oriented architecture for complex agent interaction, a core service for global common understanding (e.g. querying an ontology), a method for peer-based information propagation (e.g. gossiping), methods for balancing loads to improve scalability, and agent monitoring for creating resource awareness.

The new concepts, methods and architectures developed will be applied to production systems as one demonstrator for a very complex heterogeneous ICT system. In order to test the methods and concepts developed, the following two Scenarios are introduced:

1st scenario: Automated definition of new production processes for new customer-specific products:

In order for a facility to be able to define new production processes, it needs methods and IT-tools to derive the necessary processes from a product description. Furthermore, an evaluation metric for assessing the performance of these processes is required. However, it mostly needs to be aware of the functions of its components, their capabilities and limitations and how each component (software or hardware) can contribute to implementing and executing a process. The case study is based on the body shop assembly at "Automotive AG" of a car side door. This is a robot based body shop assembly line, where the joint operations are mainly completed by welding and gluing robots. At the beginning of the case study the side door of the so called C-Class is being produced on the line. However to decrease investment cost the Automotive AG wants to integrate the production of the so called G-Class side door in the existing C-door assembly line, instead of building a new line. This is common practice in the automotive industry. This process is very time and effort consuming as it involves several engineering & production departments. It involves numerous production-oriented product validation processes which have to be done by the engineers, as new production processes are defined for each part and assembly and validated through simulations. A self-aware production system would reallocate these tasks to the machines, robots and devices at the line, thus relieving the engineers. The self-aware production system would analyze the G-Class side door (e.g. using feature recognition) and the devices in the production system would try communicating which parts they could manufacture and/or assemble, considering their current

work-load. This would have the added value that if the devices do not succeed in defining a new production process, they will at least be able to identify the parts or assemblies they could not process thus reducing the amount of preplanning and validation to specific parts in the side door.

2nd Scenario: Resolving a disruption or malfunction within the production system:

Unforeseen changes often take place in the production domain and possibly have a negative impact on a plant's productivity and flexibility required for a production system to follow demand. In order for to maintain the material flow undisturbed, the implementation of methods and IT tools providing decentralized decision making and real time response is necessary. A centralized solution for redirection of material flow is only feasible for small production system. The present case study is based on the redirection of the material flow due to an unforeseen machine breakdown. Facility machines and products are interconnected within a real time information exchange network, utilizing a peer to peer information exchange between heterogeneous and mobile devices and sensors together with factory automation system. While a centralized solution for redirection of material flow is applicable, it is not feasible for large-scale production system. A self-aware decentralized system has the added advantages of scalability, adaptability, and robustness.

A machine or group of machines processing a product can be regarded as a decision unit. So, when a machine breakdown occurs, the agent associated with the specific machine, or group of machines, notices the problem and generates a series of alternatives for the redirection of the material flow, taking into account the product characteristics as well as the state and the characteristics of the interconnected machines. Afterwards, the impact – consequences of each alternative is determined with respect to different criteria, which have been defined according to overall facility's objectives. Finally, one or more decision-making rules are applied and the best alternative is selected. The self-aware production system, equipped with IT systems providing decentralized multi-criteria decision-making and real-time information sharing, is able to overcome unexpected errors such as machine breakdowns, without reducing the facility's performance in terms of flexibility and productivity.

4. Conclusion

This paper is focused around a long term, visionary and challenging goal: Self-Awareness within modern ICT systems and networks. Its vision is steps beyond self-organizing properties and moves towards cognitive networking and processing within systems, leading to greater flexibility, increased productivity reinforcing the EU's position in business and providing a leadership in new services.

Applying the proposed methods. models and algorithm to ICT systems encountered in the daily life, e.g. traffic lights, communication networks, transportation will improve working conditions and collaboration among employees, departments and companies, as well as improve the quality of life of the European population, especially the elderly by the seamless and ubiquitous support rendered by the self-aware system.

Acknowledgements

This work has been partially funded by the Karlsruhe Institute of Technology (KIT). The authors wish to acknowledge the KIT for its support. They also wish to acknowledge their gratitude and appreciation to Prof. J. Ovtcharova (Director of the Institute for Information Management in Engineering) for her contribution during the development of various ideas and concepts presented in this paper.

References

1. Loveless, *The Role Of It: Practical Issues For The Primary Teacher,* London: Cassell, 1996.
2. Commission, European. "Overlay Computing & Communication", European Commission, Brussels, 2008.
3. H. Sakai, and K. Amasaka, "The robot reliability design and improvement method and the advanced Toyota production system," in *Industrial Robot*, Vol. 34, No. 7, 2004:, pp.310-316.
4. C. Côté, Y. Brosseau, D. Létourneau, C. Raïevsky, and F. Michaud, "Robotic Software Integration Using MARIE" in *International Journal of Advanced Robotic Systems* - Special Issue on Software Development and Integration in Robotics, 2006, pp. 55-60.
5. W. Li, A. Joshi, and T. Finin, "Policy-based Malicious Peer Detection in Ad Hoc Networks," presented at the International Conference on Information Privacy, Security, Risk and Trust (PASSAT 2009), Vancouver, Canada, 2009, pp. 76-82.
6. H. Huang, G. Wang, and Y. Wu, "A direct approach for incomplete information systems" in Proceedings of *SPIE*, Vol. 5433, 2004, pp. 114-121.
7. D. O'Sullivan, and D. Lewis, "Semantically driven service interoperability for pervasive computing," presented at the 3rd ACM international workshop on Data engineering for wireless and mobile access. San Diego, USA, 2003, pp. 17 - 24.
8. J. Cámara, C. Canal, and G. Salaün, "Multiple Concern Adaptation for Run-time Composition in Context-Aware Systems," in *Electronic Notes in Theoretical Computer Science (ENTCS)*, vol. 215, 2008, pp. 111-130.
9. D. Jones, *The Shortcut Guide to Automating Network Management and Compliance*, Realtime Publishers, 2006.

10. R. Chadha, and L. Kant, *Policy-Driven Mobile Ad hoc Network Management*, Wiley & Sons, 2007.
11. J.P. Van Overschelde, "Requirements for a Self-Aware Computational Systems;" presented at the Darpa Workshop on Self-Aware Computer Systems, Washington DC, USA, 2004.
12. R. Raffaeli, M. Germani, and M. Mengoni, "Multi-level Representation for Supporting the Conceptual Design Phase of Modular Products," In *e Future of Product Development*, F.-L. Krause, 209-224. Heidelberg: Springer, 2007, pp. 209-224.
13. G. Pavlin, P. De Oude, and J. Nunnink, "A MAS approach to fusion of heterogeneous information," in *Proc. 2005 IEEE/WIC/ACM International Conference on Web Intelligence*, Compiègne, France, 2005, pp. 802-804.
14. A.C. BOURY-BRISSET, "Ontology-based Approach for Information Fusion," presented at the 6th International Conference on Information Fusion, Queensland, Australia, 2003, pp. 522-529.
15. J. Kludas, E. Bruno, and S. Marchand-Maillet, "Information Fusion in Multimedia Information Retrieval," in *Adaptive Multimedial Retrieval: Retrieval, User, and Semantics*, Heidelberg: Springer, 2008, pp. 147-159.
16. R. Farahbod, U. Glässer, and H. Wehn, "Dynamic resource management for adaptive distributed information fusion in large volume surveillance, " presented at the SPIE Defense & Security Symposium, Orlando, USA, 2008.
17. Y. Wang, and V. Varadharajan, "Role-based Recommendation and Trust Evaluation," in *Proc. of the IEEE Joint Conference on E-Commerce Technology (CEC'07) and Enterprise Computing, E-Commerce and E-Services (EEE'07)*, Tokyo, Japan, 2007, pp. 278-295.
18. Q. Mahmoud, *Cognitive Networks: Towards Self-Aware Networks*, West Sussex: John Wiley & Sons Ltd., 2007.
19. R.W. Thomas, L.A. DaSilva, and A.B. MacKenzie, "Cognitive networks," presented at the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, Baltimore, USA, 2005, pp. 352-360.
20. J. Tozicka, M. Rovatsos, and M. Pechoucek, "A framework for agent-based distributed machine learning and data mining," presented at the 6th International joint conference on autonomous agents and multi-agent systemes, New York, USA, 2007, pp. 1-8.
21. M. Sifalakis, M. Fry, and D. Hutchison, "A Common Architecture for Cross Layer and Network Context Awareness," In *Self-Organizing Systems: Lecture Notes in Computer Science*, Vol. 4725, Heidelberg: Springer, 2007, pp. 103-118.

Development of information-analytical technologies for the design of power-saving window systems of new generation acting as solar batteries

Rusyak I.G.
Faculty of Applied Mathematics
Izhevsk State Technical University
Izhevsk, Russia
e-mail: primat@istu.ru

Korolev S.A.
Faculty of Applied Mathematics
Izhevsk State Technical University
Izhevsk, Russia
e-mail: stkj@mail.ru

Abstract

The classification and development stages of window structures are given. The mathematical model of hydrodynamics and heat exchange in multi-layer window system is developed. The technique and programming and computing complex for the design of power-saving window systems that allow optimizing the structure parameters to obtain the required operational characteristics are demonstrated. The heat characteristics of window systems of different designs are investigated.

1. Introduction

In modern construction windows are multifunctional elements of outside protective structures of buildings. Apart from their main function – natural lighting provision, windows have to possess high heat- and sun-protective, strength and decorative properties.

Let us consider the classification and development stages of window systems.

First generation of windows – windows in wooden frames, widely applied in mass construction works in the last century. Low manufacturability and cheap materials resulted in rather cheap windows. However, such windows are non-durable, they easily lose operational properties, have low heat-insulation and air tightness, require constant maintenance and repair.

When new materials and technologies appeared, the second generation of windows became widely spread. These double-glazed windows in plastic frames are currently used. Such windows are rather durable, do not require any operational costs and possess good heat-protective properties.

The third generation of windows comprises the structures with increased operational characteristics. The application of selective low-emission glass coatings, interglass spaces filled with special gases gives the possibility to significantly increase the window heat-exchange resistance without the decrease in light transmission. These technologies also allow considerably decreasing the entry of heat energy of solar radiation through the windows in summer.

The fourth generation of windows extends the functional possibilities of light-transparent protective structures. The large area of window structures renders the possibility to use them as devices accumulating the energy of solar radiation. The application of modern selective coatings allows obtaining the required properties of light transmission and heat radiation absorption. If the interglass spaces are filled with liquids, the window can be considered as light-transparent flat solar collector. At the same time, the heat energy obtained can be used for heating and hot water supply.

There are technologies of electric solar cell production based on boron nitride (BN), transparent for visible light. The application of this technology in window systems allows obtaining the electric energy.

Thus the complex approach is necessary to designing window systems of the new generation, considering them as an important element of the sustainment in modern construction.

2. Problem definition

In this connection, the mathematical model of heat-exchange through the window system has been developed, which, in general, is considered as multi-layer inertialess system consisting of glass cladding and closed interglass interlayers (Fig. 1). Each layer has own physical and geometrical parameters defined by the glass cladding material, glass coating, filling type of interglass spaces and layer thickness.

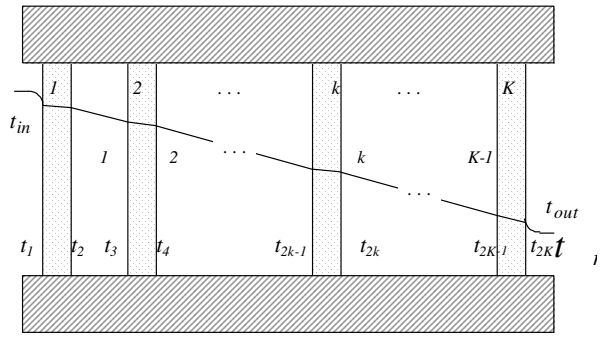


Figure 1. Window system scheme

We will consider this system as a multi-layer wall, where the layers $k = 1..K$ – glass cladding; layers $n = 1..K-1$ – interglass interlayers; t_{2k-1}, t_{2k} – temperatures on left and right surface k – glass cladding layer.

The conditions of equalities of temperatures and heat flows are assigned at the media (layers) interfaces:

$t_i(\delta_i) = t_{i+1}(0)$, $i = 1..2K-1$;

$$\alpha_{in}(t_{in} - t_1(0)) = -\lambda_1 \frac{\partial t_1(0)}{\partial x} + q_1^r(0);$$

$$-\lambda_1 \frac{\partial t_1(\delta_1)}{\partial x} + q_1^r(\delta_1) = -\lambda_2 \frac{\partial t_2(0)}{\partial x} + q_2^r(0);$$

$$\ldots$$

$$-\lambda_i \frac{\partial t_i(\delta_i)}{\partial x} + q_i^r(\delta_i) = -\lambda_{i+1} \frac{\partial t_{i+1}(0)}{\partial x} + q_{i+1}^r(0); \quad (1)$$

$$\ldots$$

$$-\lambda_{2K-2} \frac{\partial t_{2K-2}(\delta_{2K-2})}{\partial x} + q_{2K-2}^r(\delta_{2K-2}) =$$

$$= -\lambda_{2K-1} \frac{\partial t_{2K-1}(0)}{\partial x} + q_{2K-1}^r(0);$$

$$-\lambda_{2K-1} \frac{\partial t_{2K-1}(\delta_{2K-1})}{\partial x} + q_{2K-1}^r(\delta_{2K-1}) =$$

$$= \alpha_{out}(t_{2K-1}(\delta_{2K-1}) - t_{out}).$$

Here α_{in}, α_{out} – resulting heat transfer factors inside and outside of the building, respectively; t_{in}, t_{out} – temperatures inside and outside of the building, respectively; λ_i – heat conductivity factor of i – layer; δ_i – thickness of i – layer; $q_i^r(x)$ – radiant heat flux passing through i – layer.

Simulation of radiant heat-exchange in multi-layer window system is described [1].

In interglass interlayers convective heat-exchange is simulated based on Navier-Stokes equation of motion recorded for incompressible viscous continuous medium in Boussinesq approximation [2]:

$\nabla \cdot \mathbf{V} = 0$;

$$(\mathbf{V} \cdot \nabla)\mathbf{V} = -\frac{1}{\rho}\nabla p + \mathbf{g}\beta(t - t_m) + (\nabla\nu, \nabla)\mathbf{V}; \quad (2)$$

$$(\mathbf{V} \cdot \nabla)t = (\nabla \frac{\nu}{\mathrm{Pr}}, \nabla)t,$$

where $\mathbf{V}$ – velocity vector; ρ – density; p – pressure; $\mathbf{g}$ – gravity acceleration vector; β – temperature factor of medium volume expansion; t, t_m – temperature of elementary volume and average temperature in all the medium volume, respectively; ν – medium viscosity; $\mathrm{Pr} = \frac{\lambda}{\nu c \rho}$ – Prandtl number; c – medium heat capacity under constant pressure.

The system of equations (2) with boundary conditions (1) was solved numerically by control volume approach. The modified algorithm SIMPLER was realized [3]. To decrease the calculation time the effective method for solving difference equations of elliptical type proposed in [4] was applied.

3. Research results

The technique of designing multi-layer window systems that take into account the influence degree of different parameters on the structure heat characteristics and allow optimizing the structure parameters to obtain maximum resistance to heat transfer was developed based on the complex analysis of the model and recommendations of window designers [1, 5].

The designing of window system with the given operational characteristics consists of the following stages:

1. Selection of window system operational conditions (weather conditions, number of floors in the building, orientation against cardinal direction, opacity, inside microclimate parameters) and required operational characteristics (light transmission, heat insulation, sun-protection, noise insulation) taking into account the acting guidelines.
2. Designing the process of window system consists in the selection of a number of glass cladding layers and glass material, heat-reflecting and sun-protective glass coatings, gas or liquid fillers of interglass interlayers, as well as the optimization of geometric parameters of the structure to obtain optimal characteristics.
3. Visualization of design decisions, thermal calculations and analysis of operational characteristics, investigation of the dependence of window system characteristics on the changes in operational conditions.

Heat characteristics of window systems of various designs were analyzed with the help of the technique developed.

The main way to increase the heat transfer resistance of the windows of the first and second generations is the increase in the number of glass cladding layers. However this results in the increase in weight and price of the window structure. The application of different materials for glass production (fiberglass, polyvinylchloride) does not contribute to the sufficient

increase in heat transfer resistance as the glass thickness in insignificant. The heat transfer resistance of the window with double glass cladding (double-pane glass) is 0.30–0.35 $m^2 \cdot °C/W$, the window with triple glass cladding (triple-pane glass) – 0.45–0.52 $m^2 \cdot °C/W$.

Windows of the third generation allow increasing the heat transfer resistance due to heat-reflecting coatings and gas fillers without additional glass cladding layers.

The application of selective, low-emission glass coatings renders the possibility to considerably decrease the radiant component of heat-exchange through the window. Two types of heat-reflecting coatings are distinguished: solid (K-glass) and soft (I-glass). The range of radiation factors for different types of glass coatings is given in Table 1.

Table 1. Radiation factor of glass coatings

No	Coating	ε
1	Without coating	0.8-0.9
2	Solid (K-glass)	0.16-0.20
3	Soft (I-glass)	0.025-0.110

To decrease the convective heat-exchange in interglass space it is filled with special gases. The most widely spread are inert gases, such as argon (Ar) and krypton (Kr). Table 2 contains their thermal properties as compared with air.

Table 2. Thermal properties of gases

Gas	ρ, kg/m^3	λ, $W/(m^2 \cdot °C)$	c, $kJ/(kg \cdot °C)$	$\nu \cdot 10^6$, m^2/s
Air	1.25	0.024	1.006	13.9
Argon	1.78	0.016	0.522	11.7
Krypton	3.74	0.009	0.249	6.7

The combined application of low-emission glass coatings and gas fillers in double-glazed windows is the most effective. At the same time, the heat transfer resistance of a double-pane glass is 0.60–0.80 $m^2 \cdot °C/W$, triple-pane – 0.80–1.05 $m^2 \cdot °C/W$.

The windows of fourth generation allow obtaining additional heat and electric energy without decreasing the main operational characteristics. Combining gas and liquid fillers in interglass interlayers, applying selective coatings, we can obtain optimal conditions of heat insulation and removal of heat energy of solar radiation in multi-layer window system, which can be considered as a light-transparent flat solar collector. At the same time, external gas fillers serve as heat insulation, medium interlayer is filled with liquid acting as heat-transfer agent, glass with low-emission coating is an absorption layer. Efficiency coefficient of the system at solar radiation power 600 W/m^2 and temperature difference of heat-transfer agent 40 °C reaches 40%, thus giving the possibility to obtain up to 50 l of hot water from 1 m^2 of the window area per day.

4. Conclusion

The programming and computing complex to design multi-layer window systems was created based on the developed technique. It contains the information base of modern materials and technologies applied in light-transparent protective structures, design and analysis unit, simulation result visualization unit. The program complex can be used for the conceptual design of prospective window systems and analysis of heat-protective characteristics of the existing window designs.

References

1. Rusyak I.G., Korolev S.A. Peculiarities of mathematical modeling and designing of power-saving multi-layer window system. Bulletin of ISTU, 2005; iss. 4:20–25.
2. Dzhaluria J. Natural convection: heat- and mass-
3. exchange. M.: Mir, 1983, pp. 399.
4. Patankar S. Numerical methods of solving the problems of heat-exchange and dynamics of a liquid. M.: Energoatomizdat, 1984, pp. 150.
5. Ilyin V.P., Livshits I.Z. Stone symmetrized method. Journal of calculus mathematics and mathematical physics, 1994. V.34, iss. 11:1558-1566.
6. Rusyak I.G., Korolev S.A. Designing of multi-layer window systems with given heat properties. Energoeffektivnost, 2004, iss. 4:71-74.

Applied Computational System "Warm" for the Automation of Thermalphysic Calculations by Grinding

Vitaly Smirnov
Votkinsk branch of Izhevsk State Technical University
Udmurt Republic, Russian Federation
e-mail: smivial@mail.ru

Abstract

Thermalphysic calculations play the important role by working out of technological processes of grinding. There is a real necessity of creation of the applied computational system for the automation of thermalphysic calculations by grinding. The nonlinear mathematical model of a non-stationary temperature field of the workpiece is considered by grinding. For the problem decision the numerical finite-difference method is used, calculation is conducted according to the explicit scheme. The mathematical model for forecasting of an error of numerical calculation depending on sizes of steps finite-difference approximations is received. The numerical method of calculation is realized in the form of the applied computational system "Warm" created on the basis of programming languages C# and C++. The structure and operating procedure of applied computational system "Warm" is described. The example of calculation of a temperature field of the workpiece and thermokinetic curves is resulted. It is shown using the computational system "Warm" for choice of the cutting parameters and parameters of the grinding wheels for different work materials. Method of an experimental research of a temperature field of the workpiece is resulted. It is considered directions of the further development of the computational system.

1. Introduction

Thesis: thermalphysic calculations play the important role by working out of technological processes of grinding. The temperature field of the workpiece defines quality of the processed surface: heat defects, microcracks, size and a sign on residual strains, phase-structural transformations in surface layer. There is a necessity for working out of the applied computational system for the decision of following problems of thermalphysic of grinding:

1) calculation of non-stationary two-dimensional temperature field of the workpiece $\Theta(y, z, \tau)$;
2) calculation of the maximum temperature of the workpiece Θ_{max} under various conditions of grinding;
3) calculation of components of a gradient of temperature.
4) calculation of dependence $\Theta(\tau)$ for the set area of the workpiece;

Tasks in view are solved by applied computational system "Warm" developed by us. System "Warm" is the important tool in hands of the technologist and the researcher.

2. Description of the applied computational system "Warm"

Let's consider surface grinding by wheel periphery (fig. 1) of the workpiece in the size L_yxL_z.

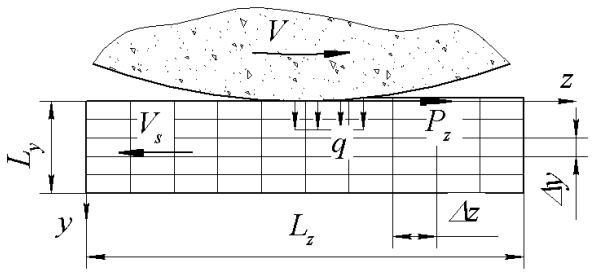


Figure 1. Solved problem

For modeling of a temperature field of the workpiece it is necessary to solve a non-stationary nonlinear two-dimensional problem of heat conductivity by grinding. This problem is described by the heat conduction equation, initial and boundary conditions:

$$\begin{cases} c(\Theta)\cdot\rho\cdot\dfrac{\partial\Theta}{\partial\tau} = \mathrm{div}\left(\lambda(\Theta)\cdot\nabla\Theta\right) \\ \Theta(y,z,0) = \Theta_0 \\ y=0: \begin{cases} \lambda(\Theta)\cdot\dfrac{\partial\Theta}{\partial y} = -\alpha_l\cdot(\Theta-\Theta_l) \quad if\ q=0 \\ \lambda(\Theta)\cdot\dfrac{\partial\Theta}{\partial y} = a_w(\Theta)\cdot q \quad if\ q>0 \end{cases} \\ y=L_y: \quad \lambda(\Theta)\cdot\dfrac{\partial\Theta}{\partial y}=0 \\ z=0,\,L_z: \quad \lambda(\Theta)\cdot\dfrac{\partial\Theta}{\partial z}=0 \end{cases} \quad (1)$$

where L_y, L_z – the sizes of processed workpiece; c - specific heat of the material of the workpiece; λ – heat

conduction coefficient of the material of the workpiece; ρ – thickness of the material of the workpiece; α_l – heat-transfer coefficient of the lubricoolant; Θ_l – temperature of the lubricoolant; q – thermal current; a_w – coefficient defining a part of the general thermal current, arriving in the workpiece

$$\alpha_w(\Theta) = \frac{1}{55{,}9 \cdot \frac{\lambda_g}{\lambda(\Theta)} \cdot \sqrt{\frac{\lambda(\Theta)}{c(\Theta) \cdot \rho \cdot d_a \cdot V}} + 1}. \quad (2)$$

where λ_g - heat conduction coefficient of the grinding wheel; d_a – average diameter of areas of deterioration of grains of an abrasive wheel; V – surface footage of the wheel.

Dependences $\lambda(\Theta)$ and $c(\Theta)$ are approximated by quadratic functions:

$$\begin{aligned} \lambda(\Theta) &= \lambda_2 \cdot \Theta^2 + \lambda_1 \cdot \Theta + \lambda_0 \\ c(\Theta) &= c_2 \cdot \Theta^2 + c_1 \cdot \Theta + c_0 \end{aligned}. \quad (3)$$

Thermal current is defined under the following formula:

$$q = \frac{P_z \cdot V}{L_c \cdot B}. \quad (4)$$

where P_z - feed force depending on characteristics of an abrasive wheel, cutting parameters, mechanical properties of the work material; B – wheel thickness; L_c – length of contact of the wheel and workpiece.

For the decision of heat conduction equation in system "Warm" the finite-difference method is used. In algorithm the explicit scheme of the decision of finite-difference equations is used [1]. For economy of computer time the temperature pays off only for the elements which are on a left side from coordinate of a right edge of a source of heat plus half of length of contact piece L_c.

For more exact representation of boundary conditions on the top border partial blanking of elements of the top layer of the workpiece by a thermal current is considered.

At realization of computational system "Warm" as the programming language has been chosen C# - the basic programming language for a platform .NET Framework. For productivity increase the basic part of calculations is carried out in the uncontrollable dll-library written on C++.

The computational system consists of the module of input of the initial data, the module of calculation of a temperature field of the workpiece, the module for calculation of components of a gradient of temperature, the module for tracking change of temperature of the set area of the workpiece, the module for calculation of the maximum temperature of the workpiece for all time of grinding, modules for adjustment of visualization of a temperature field.

The scheme of work of system "Warm" is presented on fig. 2.

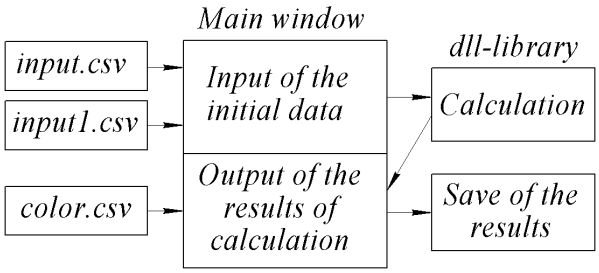


Figure 2. The scheme of work of system "Warm"

As the initial data in the computational system the sizes of the workpiece, quantity of elements along axis z and y, calculation step on time, parameters of an abrasive wheel and grinding parameters are set. Thermalphysic and mechanical characteristics of a processed material and an abrasive wheel are read out from files input.csv and input1.csv. The user can edit these files, adding new materials and abrasive wheels. The user enters number of time readout then the program directs the initial data to uncontrollable dll-library. The library makes all necessary calculations and sends results in the main program. The result of calculation can be presented in the form of the table of values or in the form of the graphic of level lines. The user can adjust color scale of the graphic of level lines. Options of color scale are stored in a file color.csv. Results of calculation of a temperature field, components of a gradient of temperature, change of temperature in a point can be kept in the form of a spreadsheet for the further processing.

The important question at use of numerical methods is accuracy of the received numerical decision. It is established that in our case the calculation error consists of two parts:

1) an error of approximation of partial derivatives in the heat conduction equation;

2) an error of instability of the explicit scheme of calculation.

Dependence of an error of the numerical decision on values of steps of calculation on coordinates and time is received:

$$\delta(\Delta y, \Delta z, \Delta\tau) = k_1 \cdot \left(\frac{1}{\Delta z^2} + \frac{1}{\Delta y^2} \right) \cdot \Delta\tau + b_1 \cdot \Delta y. \quad (5)$$

where δ – error of the numerical decision, %; k_1, k_2, b_1 – the factors depending on the processed material and speed of movement of a source of heat; Δy, Δz, $\Delta\tau$ – calculation steps on coordinates and time.

Mathematical model (5) allows choosing sizes of steps of the finite-difference approximation providing set accuracy of calculation. This possibility is realized in system "Warm".

Applied computational system "Warm" has following advantages in comparison with existing analogues:

1) the system considers distinctive features of modeling of process of grinding (grinding parameters, mathematical models of grinding force, features of interaction of an abrasive wheel with the workpiece,

change of cutting power of an abrasive wheel, presence of lubricoolant and others);
2) the system is simple in use and does not demand special training;
3) the system can be easily regulate for the decision various thermophysical problems.

3. Results got by means of system "Warm"

The example of calculation of a temperature field of the workpiece is shown on fig. 3. The initial data for calculation: surface footage of grinding wheel V=35 m/s, cutting rate V_s=5 m/min, work material – high-speed steel, material of the grinding wheel – cubic boron nitride (CBN), specific cutting power on 1 mm of wheel thickness N_{sp}=75 W/mm, sizes of area 8 x 2 mm. Maximal calculation temperature has made 550,9 ºC.

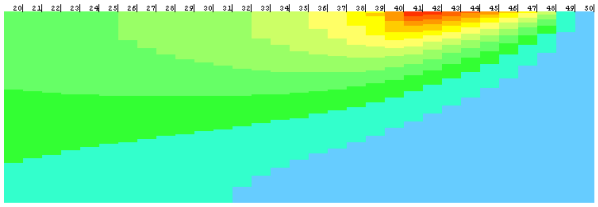


Figure 3. Example of calculation of temperature field of the workpiece

System "Warm" application allows choosing grinding parameters and the parameters of a grinding wheel providing the demanded thermal mode of grinding. On fig. 4 the graphic of dependence of the maximum temperature of the workpiece Θ_{max} from size of specific cutting power N_{sp} and cutting rate V_s for wet grinding of stainless steel X6CrNiTi18-10 by a solid grinding wheel from electrocorundum, on fig. 5 similar dependence for wet grinding of a titanic alloy Ti-6Al-4V by solid grinding wheel from silicon carbide, on fig. 6 similar dependence for dry grinding of a steel S6-5-2-5 by solid grinding wheel from CBN.

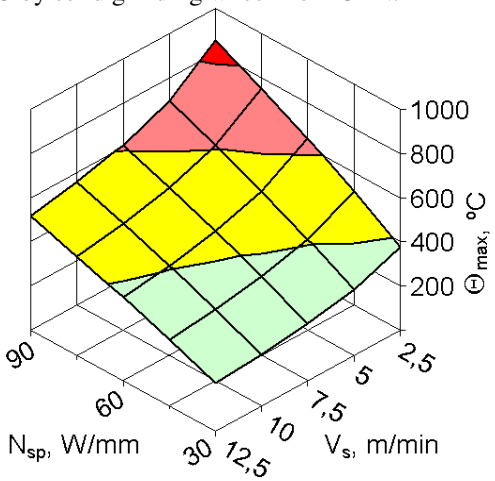


Figure 4. Dependence $\Theta_{max}(N_{sp}, V_s)$ by grinding of a stainless steel X6CrNiTi18-10 by grinding wheel from electrocorundum

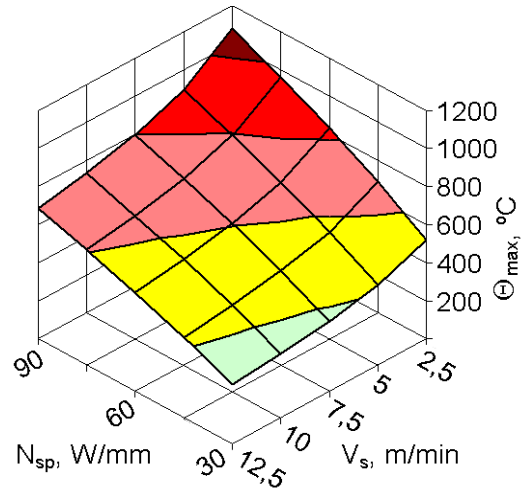


Figure 5. Dependence $\Theta_{max}(N_{sp}, V_s)$ by grinding of a titanic alloy Ti-6Al-4V by grinding wheel from silicon carbide

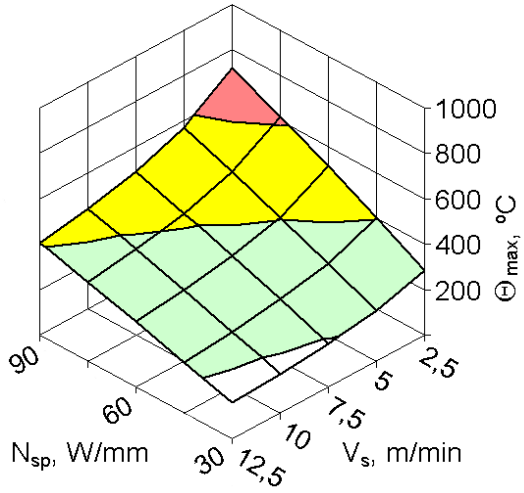


Figure 6. Dependence $\Theta_{max}(N_{sp}, V_s)$ by grinding of a steel S6-5-2-5 by grinding wheel from CBN

On fig. 7 calculation curves of changes of temperature in time $\Theta(\tau)$ on various depth of the workpiece (work material – stainless steel X6CrNiTi18-10, solid grinding wheel from electrocorundum, N_{sp}=90 W/mm, V_s=2.5 m/min).
On size of the maximum temperature and finding time at temperature above critical it is possible to judge depth of phase-structural transformations in surface layer of the workpiece.

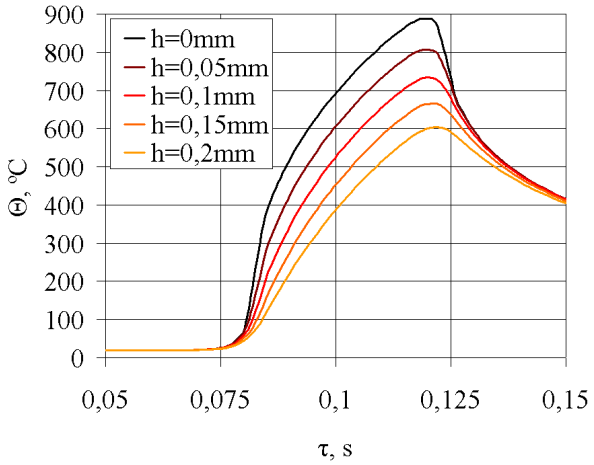


Figure 7. Thermokinetic curves on various depth of the workpiece

Applied computational system "Warm" gives the ability to choose parameters of the grinding wheels providing the set thermal mode of grinding. For example, we will consider surface grinding by wheel periphery of a titanic alloy Ti-6Al-4V. Grinding parameters: V=35 m/s, V_S=5 m/min, N_{sp}=45 W/mm. On fig. 8 thermokinetic curves on a surface of the workpiece for a solid grinding wheel, interrupted grinding wheel with length of a ledge l_1=0.025m and interrupted wheel with l_1=0.02 m are shown.

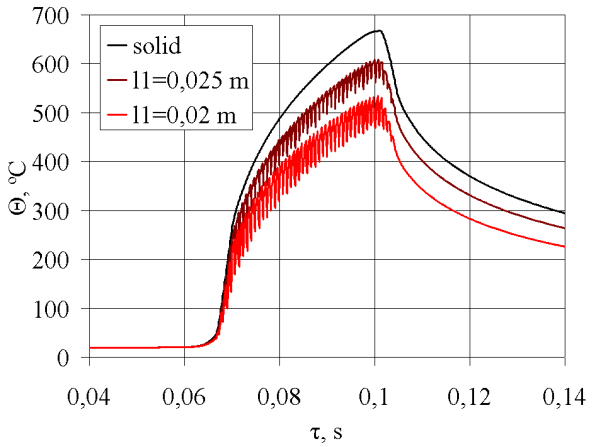


Figure 8. Thermokinetic curves for solid and interrupted grinding wheels

Fig. 8 shows that use of interrupted grinding wheels allows lowering the maximum temperature of the workpiece at constant specific power of grinding.

For experimental check of the received results the method of the semiartificial thermocouple is used (fig. 9). One of thermocouple electrodes is the processed material 1. At grinding arises thermojunction therefore arises thermocouple voltage. Signal registration is carried out by means of the analog-to-digital coder LTR27 which is a part of crate system LTR U-8-1 [2] – position 3 on fig. 9. The obtained data is transferred in the computer 4. For mathematical processing of experimental data specialized software ACTest is used.

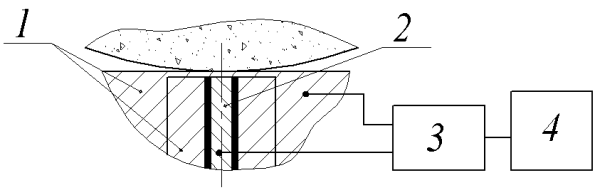


Figure 9. The scheme of measurement of temperature

Experimental researches have shown that the error of calculation of a temperature field of the workpiece by means of applied computational system "Warm" does not exceed 3 %.

4. Conclusion

The developed computational system is the effective tool for working out of technological process of grinding. Use of similar system allows exercising administration of process of grinding of any material at the expense of a choice of parameters of process.

Directions of the further development of computational system "Warm":

1) creation of the module of forecasting of structure in surface layer after grinding on the basis of the analysis of change of temperature on various depth from a surface;

2) creation of the module of forecasting of temperature strains in surface layer of the workpiece.

References

1. Turchak L.I., Plotnikov P.V. Bases of numerical methods, Moscow, 2003.
2. http://www.lcard.ru/products/ltr/ltr - CRATE SYSTEM LTR - system of data gathering and steering with interfaces USB 2.0 and Ethernet (TCP/IP).

Distributed Telerobotic Control and Adaptation of Autonomous Mobile Robots

Sergey Moskvichev
Technical Cybernetics Dept.
Ufa State Aviation Technical University
Ufa, Russia
e-mail: mosk.sergey@gmail.com

Rustem Munasypov
Technical Cybernetics Dept.
Ufa State Aviation Technical University
Ufa, Russia
e-mail: rust40@mail.ru

Abstract[1]

The Internet-based telerobotics becomes more viable today due to increasing connection speeds, and allows for distributing various robotic tasks among operators all over the world. In this research-in-progress we present a theoretical study on a telerobotic system that enables remote operators to take part in control and adaptation of mobile robots through the Internet. The operators controlling robots perform various robotic tasks while watching the process through a live video stream with a supporting augmented reality layer, and their actions can then be used for learning the controller system of the robots, presenting it with multiple strategies of performing the robotic tasks.

1. Introduction

Distributed computing and Internet technologies have opened new application perspectives for robot teleoperation systems. Examples of such telerobotic systems are tele-teaching/tele-learning, virtual laboratories, remote and on-line equipment maintenance, and projects requiring collaboration of remote experts. Since teleoperated mobile robots are used in a broad range of fields, from human-assistance to military applications, it seems natural to expect that many of these applications will be (or already are) influenced by the advantages provided by the spread of the Internet and its increasing bandwidth.
One of the first successful World Wide Web based robotic projects was the Mercury project [1]. This later evolved in the Telegarden project, which used a similar system of SCARA manipulator to move objects within a defined workspace. Users were able to control the position of the robot arm and view the scene as a series of periodically updated static images. Video streaming technology is much less limiting and enables more precise control of the situation. Since the transfer of fluent video via the Internet requires a high bandwidth capacity, different methods of video transmission, such as Real Video, QuickTime Movie, Microsoft Media Player and Adobe Flash, are used [2].

Making human-robot interaction more natural and efficient is crucial when performing tasks that need as complete control of the situation as possible [3]. In particular, grounding, situational awareness, a common frame of reference and spatial referencing are vital in effective communication and collaboration. Augmented Reality (AR), the overlaying of computer graphics onto the real worldview, can provide the necessary means for a human-robotic system to fulfill these requirements for effective collaboration. Human-robot interactive systems can significantly benefit from AR technology because it conveys visual cues that enhance communication and grounding, enabling the human operator to have a better understanding of what the robot is doing and its tasks.
Many enterprises experience difficulty in training people to work with expensive equipment, which is needed for carrying out profitable work tasks (e.g. production line robots). Similar problems are found when work is of a complex and safety critical nature (e.g. nuclear environments, explosive placement, surgery) [2]. AR-based technology of telepresence provides an elegant and cost-effective solution to this problem, as it becomes more utilizable with the increasing bandwidth of computer networks and lowering computational cost of real-time video processing.

2. Telerobotic System Architecture

In order to build a successful setup for remote experiments, a number of requirements have to be fulfilled [4]:

- Telecontrol

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

- Telepresence
- Data collection
- Scheduling
- Security
- Logging
- Synchronous communication
- Collaborative environment

The system described here is a teleoperated robotic laboratory with two LEGO Mindstorms NXT mobile robots performing a robotic task. Each robot can be remotely controlled through the Internet by a remote operator with a Java-based client program. The robots share a competitive task of collecting virtual objects represented by color markers in the AR layer. The operators can view the scene through a fixed web-camera, and the video stream is augmented with the AR layer, which carries important information about the robots and the scene, such as the virtual markers, battery levels and information about the current operators. The scene can also be viewed by others through a conventional web browser supporting Flash (fig. 1).

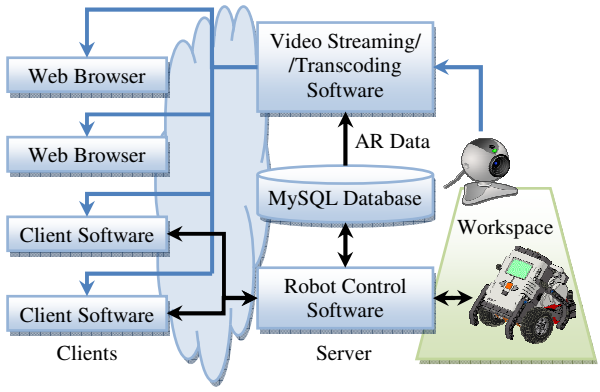


Figure 1. Telerobotic System Architecture

3. Robotic Learning

Since the robots teleoperated by multiple users are "experiencing" various control strategies applied by their operators while performing the task, this diversity can be utilized to learn the robots to complete their tasks.

The robotic control system presented here extracts the information from the environment, including the distances and directions to the virtual marker and the other robot and velocity vectors of the robots, and stores it in a quantized form to reduce the number of dimensionality of the learning problem. After each significant change in the operator's activity (like starting of stopping movement, abrupt turning) the environment state is monitored. According to the modified episode-based reinforcement learning method (EBRL), a recall table is used to keep a number of pointers to memory elements which seem similar to the current situation [5]. Each element of the table is a pair of pointer and weight which indicates similarity between the memorized episode and the current situation. The learner replaces each element by the pointer of next memory element following along the forward link in the memorized episode, and modifies the weight by combination with the similarity between the sensory data in new element and current sensory input by

$$\Delta W_i = \alpha \cdot (S(M_i, s) - W_i), \qquad (1)$$

where W_i is the weight of the ith element, $S(M_i,s)$ is similarity between the current input s and the memorized data M_i pointed from the ith element, and α is a constant of (0,1). In addition to renewed elements, the other memory elements similar to the current sensory input are also candidates of recall table elements. Assigning the value of similarity as the weight of newly recalled memory element, N_r elements of the candidates with relatively larger weights remain in the recall table, where N_r is the capacity of the table.

Similarity $S(x,y)$ between vectors x and y is calculated by

$$S(x, y) = 1 - \sqrt{\frac{1}{2n}\sum_{i=1}^{n}\frac{(x_i - y_i)^2}{V_i}}, \qquad (2)$$

where n is a number of elements in a vector, x_i and y_i are ith elements of x and y, V_i is a variance of the ith element among all of the memorized data.

The agent simply memorizes the experience as described above, and performs actions which most resemble the current situation. This approach allows utilizing the numerous performances provided by remote operators in a time- and computationally-efficient way which is important for online real-time operation.

4. Conclusion

The Internet-based distributed telerobotic system presented in this research enables remote teleoperation of mobile robots in a laboratory environment for multiple users across the world. Augmented reality live video streaming allows for consistent telepresence operation, and the diversity of control strategies presented by the operators can then be used to learn the robots to perform their tasks.

The next steps in this research-in-progress are: comparison and selection of video streaming and data transfer protocols; comparison of various controller architectures and learning algorithms, including artificial neural networks; hardware and software implementation of the system; tests and experiments performed in a real environment.

References

1. Goldberg K., Maschna M., Gentner S., et al. “Desktop Teleoperation Via The WWW”. In: *Proc. of the IEEE International Conference on Robotics and Automation*, 1995, pp. 654-659.
2. Safaric R., Debevc M., Parkin R. M., Uran S. “Telerobotic Experiments via Internet”. *IEEE Transactions on Industrial Electronics*. 2001; 48(2): 424-431.
3. Green S. A., Billinghurst M., Chen X., Chase G. J. “Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design”. *International Journal of Advanced Robotic Systems.* 2008; 5(1): 1-18.
4. Hoyer H., Jochheim A., Rohrig C., Bischoff A. “A Multiuser Virtual-Reality Environment for a Tele-Operated Laboratory”. *IEEE Transactions on Education.* 2004; 47(1): 121-126.
5. Unemi T. “Scaling up Reinforcement Learning with Human Knowledge as an Intrinsic Behavior”. In: *Proc. of the 6th International Conference on Intelligent Autonomous Systems*, 2000, pp. 511-518.

Formation of Structure of Information System on the Basis of Monitoring Fuzzy Relations of Business – Processes

V.V. Antonov
Ufa state aviation technical university
Ufa, Russia
e-mail: boss@bashkortostan.ru

G.G. Kulikov
Ufa state aviation technical university
Ufa, Russia
e-mail: gennadyg_98yahoo.com

Abstract[1]

Modelling of information system has the general philosophical basis. The philosophical concept most essential to modelling is the subject domain which can be determined as mentally limited area of reality or the area of ideal performances subject to modelling, consisting of objects taking place in the certain relations among themselves and having different properties. Thus property is understood as prominent feature of object for which rating the certain measure - a parameter of property which is characterized by set of values is established**.**

1. Introduction

The conventional formal definition of concept of a subject domain till now is absent. Definition of a subject domain as parts of the real world or set of classes of real objects subject to modelling, assumes modelling reflection with the purpose of studying under the certain corner of sight. This corner of sight itself enters into concept of a subject domain. Last years in the theory of databases and information storehouses the new directions of the scientific researches which have received the name of the theory of conceptual modelling in databases, corporate storehouses of the data were generated. The subsequent development of the created theory became development of methods of modelling of subject domains in information systems. At research of a subject domain the significant amount of the information which has subjective character can be received. Its performance in natural language contains illegibilities or uncertainty which have no analogues in language of traditional mathematics. At definition of concept of a subject domain it is necessary to take into account the following methodological aspect, by consideration of a subject domain as parts of the validity - ontologic, and by consideration of a subject domain as knowledge about this validity - gnosiological. And as result we have two various classes of models and a problem of search of conformity of the given models of the validity.

2. Problems of construction of formal model of the subject domain

Translation of conditions of a practical problem on language of mathematical models always was difficult and frequently resulted in loss of difficultly - formalized qualitative information. Many modern problems of management simply may not be solved by classical methods because of very big complexity of mathematical models, their describing. In conditions of application of the automated systems there is a transformation of functions of the person, there are new communications between the person and system, some functions are completely transferred system. The computerization promotes expansion of opportunities of the subject, sometimes qualitatively changing the maintenance of his activity. And this activity should be precisely determined and fixed for the subject the same as and behind the automated system.

Process begins with allocation of objects of a subject domain and revealing of communications between them. We shall consider more in detail process of a choice of objects. In a general view function of a choice can be presented as set of the alternative objects chosen on some condition which in turn may to be submitted as set of data on a condition of object and set of rules of a choice.

Let σ - set of data describing objects, Z - set of objects of a subject domain, $z_i \in Z$ - object from set of objects. It is obvious, that the part of data describing object can be presented as set of its information characteristics $x_i = \{\langle A_i, D_i \rangle \mid x_i \subset \sigma\}$, where A_i - nonempty set of names of properties (attributes) object with number i, D_i - set of values of the appropriate attributes, x_i - set of information characteristics object with number i.

The dictionary of elements of allowable values subdivided on classes that allows to present a subject domain as hierarchical structure may be made. Values are broken into classes of objects which cooperate with each other on the basis of rules. Let π - set of rules of a choice then conditions of a choice of object of set of alternatives may be submitted as a train $y = \langle \sigma, \pi \rangle$. On set of attributes relations $G = \{\overline{G}, \tilde{G}\}$ which share on

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

quantitative $\bar{G}$ and qualitative $\tilde{G}$ may be established and the set of types of a choice, for example $T = \{\text{"conformity","equivalence","preference"}\}$ is determined. Then any rule of a choice may be submitted by a train $\pi = \langle G, T \rangle$.

Thus, the information on object $z_i \in Z$ can be presented as set of information characteristics of the object, the established relations and rules of an establishment of relations

$$z_i = \{x_i, G, \pi\} = \{\langle A_i, D_i \rangle, \bar{G}, \tilde{G}, T\}, i \in N.$$

Values of attributes may have not numerical character. In particular, in macroeconomic, sociological, marketing, medical, legal storehouses of the data the linguistic form of data presentation is widely used.

Generally the characteristic of each object x_i may be described by the appropriate linguistic variable $\langle A_j, T_j, D_j \rangle$, where $T_j = \{T_1^j, T_2^j, ..., T_{m_j}^j\}$ - term - set of a linguistic variable A_j (a set of linguistic values of attribute), m_j - number of values of attribute; D_j - (a subject scale) base set of attribute A_j. For the description of terms $T_k^j, k = 1, ..., m_j$ appropriate to values of attribute A_j, fuzzy variables $\langle T_k^j, D_j, \tilde{C}_k^j \rangle$, i.e. value T_k^j may be used - is described by fuzzy set $\tilde{C}_k^j = \{\langle \mu_{C_k^j}(d) \mid d \rangle\}$, $d \in D_j, k = 1, ..., m_j$ in base set D_j. Then as the fuzzy characteristic of object x_i the fuzzy set of the second level may be taken

$$\tilde{x}_i = \{\langle \mu_{x_i}(a_j) \mid a_j \rangle\}, \mu_{x_i}(a_j) = \bigcup_{k=1}^{m_j} \{\langle \mu_{\mu_{x_i}}(T_k^j) \mid T_k^j \rangle\}, \quad T_k^j \in T_j, a_j \in A_i \qquad (1)$$

Proceeding from given, the subject domain can be presented as the multilevel environment consisting of set of elements of a subject domain, set of functions and the methods working on these elements and set of properties of elements and relations between elements, i.e. as ontology which includes the description of properties of a subject domain and interaction of objects in some formal language having logic semantics. If system difficult, the number of factors is great, the account of all of its characteristics results in extreme complexity. Therefore it is necessary to enter only limited number into model, and the staying components to take into account obviously not entering in model, but taking into account their influence as fuzzy reaction of model to this or that choice of alternative. It is obvious, that algebraic comparison a component is impossible and may be executed with application of methods of fuzzy logic.

Each object of a subject domain is described by values of parameters. If the set of objects of a subject domain has the common properties, or values of attributes their describing, we may speak about presence of a class of objects of a subject domain. Under procedure of formation of classes, we shall understand ordering objects on their similarity a method clasterization. The problem consists in splitting objects on clasters when as a basis of splitting the vector of parameters of object serves.

Under clasterization we shall understand representation of objects as a vector of the numbers determining a geometrical arrangement of object in space which coordinates are attributes of object. Group of the objects forming in space compact somewhat area, we shall name claster. At work with difficult objects clasters may be blocked, have dim borders that occurs as a result of the incomplete or fuzzy information on object.

Thus, the final set of objects $Z = \{z_1, ..., z_n\}$ of a subject domain may be used as set of objects clasterization. The given set is described by final set of attributes, each of which quantitatively represents some property or the characteristic of elements of a considered subject domain. Use of means of fuzzy logic allows to approve that for each object in some quantitative scale all values of attributes are measured, that is some vector in which coordinates is put to each object in conformity are quantitative values of the appropriate attributes. We shall characterize the objects $Z = \{z_1, ..., z_n\}$ subject clasterization by vectors of parameters (attributes) $y^i = \{y_1^i, ..., y_{p_i}^i\}$, where p_i - number of attributes describing object with number *i* object. Now one of the most known algorithms of realization fuzzy clasterization is the algorithm of Fuzzy c-means. Thus the basic difference of methods fuzzy clasterization from precise is introduction for each object of functions of an accessory to various clasters.

Let $F = \{F_1, ..., F_k\}$ - set fuzzy clasters a subject domain, $\mu^F = \{(\mu_1^1, ..., \mu_k^1), ..., (\mu_1^n, ..., \mu_k^n)\}$ - set of functions of an accessory of objects fuzzy clasters. Then the sum of degrees of an accessory to all clasters for each object i satisfies to a condition $\sum_{j=1}^{k} \mu_j^i = 1$.

We shall enter the additional restriction, based that fuzzy clasters form an fuzzy covering of set of objects in that and only in the event that for everyone claster j it is carried out $\sum_{i=1}^{n} \mu_j^i \geq 1$. After the task of degrees of the accessory satisfying the specified restrictions, we may calculate coordinates of the centres claster v^j under the formula

$$v^j = \frac{\sum_{i=1}^{n}(\mu_j^i(y^i)^2 \times \sum_{m=1}^{p_i} y_m^i)}{\sum_{i=1}^{n}\mu_j^i(y^i)^2},\ j=1,...,k,$$

(2)

The problem is reduced to minimization of criterion function (5) at the specified restrictions.

$$J(Z) = \sum_{m=1}^{k}(\sum_{i=1}^{n}(\mu_j^i(y^i)^2 \times \sum_{m=1}^{p_i}(y_m^i - v^j)^2))$$

(3)

Let in process clasterization the fuzzy covering is constructed. By consideration of the next object on conformity to everyone claster $F = \{F_1,...,F_k\}$ and change of the appropriate functions of an accessory, at occurrence of a situation when $\sum_{j=1}^{k}\mu_j^{n+1} < 1$, it is made a decision on creation new fuzzy claster. At absence of the information on what values or attributes of a vector of parameters the degree of an illegibility of an accessory of object to claster is defined by the maximal distance from the centre claster. Understanding as the alternative decision variant of a choice taken into account a component, we shall designate through X set of alternative decisions. The choice of the most preferable decision, in each concrete case, may be carried out on set of difficult criteria with normalization its component. On the basis of the given set X we shall generate set of the ordered pairs $E = X \times X$ alternative decisions. We shall designate through (x,y), where $x \in X$, $y \in X$ a pair of alternative decisions. Having designated through $\mu(x,y) \in [0,1]$ - function of an accessory of the fuzzy relation of preference of the decision x before the decision y. The fuzzy relation of preference we may define the formula $P = \{E, \mu(x,y)\}$. It is obvious, that for everyone P there is a return fuzzy relation $P^{-1} = \{E, \mu(y,x)\}$. We shall define a degree of the superiority of the decision x above the decision y as $\Delta(x,y) = \mu(x,y) - \mu(y,x)$. It is obvious, that $\Delta(x,y) = -\Delta(y,x)$. Then the set of not dominant decisions may be determined $X^{und} = \{x \mid \Delta(y,x) \le 0, \forall y \in X\}$. Taking into account, that us interest only set precisely nondominated decisions, having excluded the relation of an equivalence $\Delta(x,y) = 0$, and using proofs given in [6], the decision is reduced to a choice of alternatives on a basis β-level relations of preference. Generally, having m pairs alternative decisions, we may enter β-level relations preference for each pair alternative decisions as:

$$F_i(\beta) = \left\{ \begin{matrix} (x,y) \mid \Delta_i(x,y) \ge \beta, \beta \in [0,1], \\ x \in X,\ y \in X,\ i = 1,...,m \end{matrix} \right\}.$$

(4)

Using conclusions given in [6], we may enter one fuzzy relation) of preference equal $\sum_{i=1}^{m}\lambda_i\mu_i(x,y)$, where λ_i - weight factors of importance for fuzzy relations of preference. Thus the decision is reduced to multicrtiterion to a problem of acceptance of decisions, as criteria of efficiency functions of an accessory and finally to calculation of one function on formal algorithm of the decision β-level models of problems of acceptance of decisions act:

$$F(\beta) = \left\{ (x,y) \mid \sum_{i=1}^{m}\lambda_i\Delta_i(x,y) \ge \beta \right\}. \quad (5)$$

The model covering information system, may be submitted as metabase which contains the information by each kind of object of the account. On the other hand information system may be submitted as functional system - i.e. as set of functions. Thus, the purposes and restrictions may be given as fuzzy sets. The interrelation between them may be determined by the relation on the Cartesian product [1]. Considering the purposes and restrictions as symmetric elements of the logic circuit, simply enough we may generate on their basis the decision which is in essence a choice of one or several of available alternatives. Thus the fuzzy decision may be considered as some "instruction" which illegibility is consequence of discrepancy of the formulation of objects in view and restrictions, i.e. influence of the fuzzy purpose and fuzzy restriction on a choice of alternatives is characterized by their crossing which forms fuzzy set of decisions**.** The case when and restrictions are set to the purpose as fuzzy sets in different spaces, may be shown to a case when they are set in the same space [1].

Business - processes considered as the functional model of real processes, represents the structured description of the given sequence carried out business - operations, that is horizontal hierarchy of functional actions internal and dependent among themselves. The treatment of business - processes as consumers and converters of some resources allows to consider them as objects of optimum control. Thus the business - process considered as set of consistently carried out chains of operations, may be treated as set of cooperating subsystems, i.e. as discrete dynamic system changed in space and time.

Modelling of information system has the common philosophical basis. The philosophical concept most essential to modelling is the subject domain which may be determined as mentally limited area of reality or the area of ideal representations subject to modelling, consisting of objects taking place in the certain relations among themselves and having different properties. Thus property is understood as prominent feature of object for which estimation the certain

measure - a parameter of property which is characterized by set of values is established**.** These values are designated by symbols from some beforehand certain set named the alphabet. Thus, property of object is a reality, and a parameter - a subjective measure of this reality. Set of parameters of properties of object is parameters of object. The conventional formal definition of concept of a subject domain till now is absent. Definition of a subject domain as parts of the real world or set of classes of real objects subject to modelling, assumes modelling reflection with the purpose of studying under the certain corner of sight. This corner of sight itself enters into concept of a subject domain. Therefore it is accepted to count the majority of researchers, that the concept of a subject domain may not be formalized as initial concept. Last years in the theory of databases and information storehouses the new directions of the scientific researches which have received the name of the theory of conceptual modelling in databases, corporate storehouses of the data were generated. The subsequent development of the created theory became development of methods of modelling of subject domains in information systems. Model of a subject domain are our knowledge of a subject domain. At research of a subject domain the significant amount of the information which has subjective character may be received. Its representation in natural language contains illegibilities or uncertainty which have no analogues in language of traditional mathematics. At definition of concept of a subject domain it is necessary to take into account the following methodological aspect, by consideration of a subject domain as parts of the validity - ontologic, and by consideration of a subject domain as knowledge about this validity - gnosiological. And as result we have two various classes of models and a problem of search of conformity of the given models of the validity:

- model of the validity;
- model of knowledge about this validity.

It agrees given to research organization Gartner Group almost half of projects on realization of information system it appears unsuccessful. The basic problem consists in not formalized process of modelling of a subject domain that does impossible application of mathematical methods of the analysis of properties of models of a subject domain, such as functional completeness and integrity. All this puts a question on consideration of a problem of modelling of a subject domain from positions of the methods which are taking into account an illegibility or uncertainty of descriptions of model of researched object. Any phenomenon represents known unity of quality and quantity. This unity is fixed in philosophy by concept of a measure of object. The space and time in themselves not is a matter, and only the ways of its life expressing that empirical fact, that all things, events and processes are in a situation subordination coordination [2]. Existence in time means the order of their following one after another, cyclicity and the certain rhythmics of change, and distinctive attributes of space and time as attributes of a matter are their extent and duration. To speak about a subject domain it is meaningful, if it has the certain semantic localization, for example in space and time or functional. Then construction of semantic model is reduced to formalization of logic relations. There are various methods of acceptance of the decisions, representing various ordering of considered decisions based on the same expert estimations. The modern condition of researches in the field of fuzzy logic testifies to increase of value of integration of fuzzy logic with other fields of knowledge. One of the basic complexities by development of model of a subject domain is, that the number of possible variants of formalization of a subject domain is indefinite. The model adequately should display any subset (variant of formalization of a subject domain), and process of modelling may have any idea, allowing to define value of any object of a subject domain by realization of any certain sequence of actions. One of such ideas - the method of syntactically-guided translation based at works Noel Chomsky. From hypothesis Chomsky [3,5] follows, that the semantic analysis may be shown to syntactic and consists of two steps:

- recognition of structure;
- construction of target actions on the basis of this structure.

Thus, the mathematical approach [3] allows to be limited to set of chains which can be defined in some exact image. That is we may speak about some formal language given as set. For construction it is necessary to have algorithm which on the given grammar builds a conclusion produced by this grammar. According to [3] such algorithms for any grammar Chomsky does not exist. Probably some ways of the decision of the given problem, consisting or in development of algorithm of recognition for each special case, or imposing of restrictions on rules of grammar, allocation of subclasses of grammar for which the algorithm exists.

Allocating objects in a subject domain, we receive, that each object is defined by final set of attributes, each attribute has set of allowable values, and we have attributive broadcast grammar according to which rules of calculation of attributes are defined and the algorithm of an attributive tree of a conclusion may be constructed. D. Knut formalized similar ideas, having entered concept of «attributive translation» [3]. To symbols of grammar the final number semantical the certain attributes which values may be numerical, symbolical, by matrixes either with what or attributes is put in conformity. Rules of grammar define communications of these attributes, and calculation of values is defined by a tree of a conclusion. All this allows to define semantics of language so that value of any attribute, in any unit of a tree of a conclusion might depend on all tree arbitrarily. From here any function which is determined on a tree of a conclusion follows,

that, may be submitted as attribute of any certain unit. A subject domain it is possible decomposition on elementary objects, each of which is described by set of attributes. Objects of a subject domain are connected among themselves certain attitudes which in aggregate can be presented as weighed on to edges in part focused the column. The structure the column represents structure of a subject domain. Subgraphs the column represent complex objects or subsystems of a subject domain. Instead of graphs for representation of structure of a subject domain it is possible to use language of the theory of sets and lattices of their splittings. Each train of a database is the description of a condition of some elementary object. The subset of all trains similar to the given train concerning a chosen measure of similarity, is representation of elementary object.

Let some object is allocated and designated by the term. As a rule, the new object is compared to already known objects, and its information model is formed as set of comparison of information models before known objects. Thus the model of a new subject domain for the given object will be under construction on the basis of a subject domain of that object which became known to the first. As a result of knowledge of a subject domain including the given object, will be structured as set of properties of the first allocated object and a sequence of changes of subject domains of the subsequent objects. It turns out, that earlier allocated objects appear in more exclusive position in relation to the subsequent objects of a subject domain as models of these objects are designed by means of change of model of already known objects. Thus the sequence of a choice of objects is will of a case. At attempt of a choice of object which in the best way approaches as the first it is found out, that it is the most convenient to use an image of some idealized, average object of this subject domain which model is replaced with a set of the variables describing objects of a subject domain in its quality. The choice of a set of variables for the description of objects of a subject domain and a choice of allowable values for these variables substantially is any. However, this choice further will define borders of applicability of its model.

Let W a researched subject domain of objects $W=\{w_1,...,w_n\}$. Frequently at models of reasonings there are fuzzy concepts, however any information which has been written down in any formalized kind and submitted in memory of the computer is precise. Therefore the illegibility of knowledge or relations may be defined by semantics of the information. The properties not included in allocated subject domains can be considered as a separate subject domain with special properties - an environment. Thus, any subject domain may be considered distributed since it cooperates with an environment. It agrees above told $w_{i+1} = F_i^W(w_i)$ where $F_i^W(w_i)$ is not functions in the usual sense, and defines only possible conditions of a subject domain of one object on the basis of difference from another. I.e. in this case the question is about possible fuzzy communications between metaobjects.

The model of a subject domain is defined by means of function of representation and family of modelling functions. Let S - model of a subject domain of objects $F^M : W \to S$. The model (may be put to each object of a subject domain in conformity on the basis of function of modelling F^M) $r_i = F_i^M(w_i)$. Thus to each object $w_i \in W$ corresponds $r_i \in S$. Thus there should be a function $F_i^R(r_i)$ which unequivocally defines r_{i+1} on r_i, i.e. $r_{i+1} = F_i^R(r_i)$. It is obvious, that r_{i+1} may be determined on a chain $w_{i+1} = F_i^W(w_i)$ and $r_{i+1} = F_{i+1}^M(w_{i+1})$. After the made substitution it is received $r_{i+1} = F_{i+1}^M(F_i^W(w_i))$. On the other hand there is, at least, one more chain $r_i = F_i^M(w_i)$ and $r_{i+1} = F_i^R(r_i)$ according to which, after the carried out substitutions it is received $r_{i+1} = F_i^R(F_i^M(w_i))$.

Comparing the received results, we may draw a conclusion, that irrespective of, whether operation in a subject domain is executed all over again, and then display to model of a subject domain is made, or display to model all over again is made, and then in model of a subject domain the appropriate operation is executed, the result will be identical. Hence, for each fixed value i it is received homomorphy F_i^W and F_i^R. The table of the variables used for the description of objects of a subject domain, usually exists in an implicit kind as the conventional set of characteristics of objects of a subject domain. It does not correspond to any concrete object from a subject domain, but only to all subject domain as a whole. To receive from it model of concrete object, it is necessary to fill in this table concrete values of variables.

Thus, the stage of construction of semantic model of a subject domain can be defined as formalization of the logic relations fixed in the description.

It is possible to imagine the description of a subject domain on the basis of the table of variables as a n-dimensional cube which each measurement corresponds to one of variables. Redundancy of the given cube since there are not all objects which may be described with the help of a set of the variables chosen for the description of objects of a subject domain is obvious. Thus, it is possible to describe and the restrictions imposed on these objects. All this may be submitted by system of the equations - formalization of model of a subject domain where crossing sets of semantic properties takes place, subject domains of objects, in connection with assumed state of distribution.

The model of a subject domain S may be submitted as metabase which contains the information on each

element of structure. On the other hand the subject domain W also may be submitted as set.

Let's enter a designation: P - set of properties determined by communications between elements of the mentioned above sets. Then the interrelation between them may be determined by the relation on the Cartesian product

$$P \times S \times W = \left\{ \begin{matrix} (p_i, r_i, w_i) : p_i \in P, r_i \in S, \\ w_i \in W, i = 1,...,n \end{matrix} \right\}. \quad (6)$$

Accessory of an element $z_i = (p_i, r_i, w_i)$ where $p_i \in P, r_i \in S, w_i \in W, i = 1,...,n$ to the given relation it is interpreted as follows: « the object of model of a subject domain r_i contains the information on property p_i of object of a subject domain w_i ».

Information search of the model of a subject domain r_i appropriate to a concrete element in object of a subject domain w_i. It is reduced to definition of the relation $R \subseteq S \times W$. Thus, it is possible to tell about any pair $(r_i, w_i) \in R : r_i \in S, w_i \in W, i = 1,...,n$, that w_i is relevant r_i and the decision of a problem of definition of relevance of elements of sets S and W, is reduced to definition of the relation $R \subseteq S \times W$.

Thus for anyone $r_i \in S$ $w_i \in W$ $r_j \in S$ $w_j \in W$, $i, j = 1,...,n$ truly, that if $w_i \subseteq w_j$ and $r_i \subseteq r_j$ that is, all elements of object of model of a subject domain r_i contain in object of model of a subject domain r_j and all elements of object of a subject domain w_i contain in object of a subject domain w_j and $(r_i, w_i) \in R$ it is carried out $(r_j, w_j) \in R$. Except for an extreme case when the relation R is the Cartesian product $S \times W$, the relation includes not all possible trains from the Cartesian product. It means, that for each relation there is a criterion, allowing to define, what trains are included into the relation and what are not present. Thus, each relation can put in conformity some logic expression P (a predicate of the relation R) dependent on the certain number of parameters (n a local predicate) and determining, whether the train (r_j, w_j) will belong to the relation R (the accessory of a train to the relation is equivalent to the validity of a predicate):

$$(r_j, w_j) \in R \Leftrightarrow \{P\} = \{S, W, R\}. \quad (7)$$

As circuits of information objects of considered model independent business - processes are determined. Processing and representation of the information, with reference to stages of considered process for quantity indicators of attributes, may be made by application of a direct method for one expert of construction of function of an accessory of the fuzzy set, for example offered Charles E. Osgood a method of semantic differentials on the basis of set of estimations on scales [2,4]. Taking into account construction of model as objects, each independent business - process we may consider as the separate (detached) part of information system. Any information interaction between difficult systems is realized consistently at physical, syntactic and semantic levels of interaction. Taking into account, that our system is divided into the information objects incorporated by semantic rules of interaction, it is possible to declare relative completeness of set of taken into account relations between elements of system which define its behaviour and are a subject of the analysis of functional reliability. Thus, relations between cooperating business - processes may be classified on the basis of mathematical rules precise and fuzzy logic. The concept of construction of such system reflects actually modern strategy so-called CALS technologies and may be considered as the tool of increase of efficiency and quality since fully complies with spirit and principles of the international standards of series ISO-9000.

3. Conclusion

The offered method allows, stage by stage allocating objects of a subject domain, establishing communications on a basis semantical the certain attributes to build model of a subject domain, reducing uncertainty of knowledge at each stage. The analysis is made at the top level of abstraction where as information objects of considered model independent business - processes of semantic blocks are determined that defines an opportunity of use of separate methods and algorithms by development and research of the broad audience of information problems.

References

1. Beniaminov E. “Algebraic methods in the theory of databases and representation of knowledge”. The scientific world, Moscow, 2003, 184 p
2. Borisov A., Krumberg O., Fedorov I. “Acceptance of decisions on the basis of fuzzy models”. Zinatne, Riga, Latvia, 1990, 184 p.
3. Knut, D. “Art of programming for the COMPUTER. Vol. 1 // The World, Moscow: 1976. 736 p.
4. Kulikov,G.G., Antonov V.V. “A method of formation of structure of storehouse of the data for the automated registration system on the basis of the process analysis of a subject domain”. Bulletin UGATU: the collection of proceedings. Ufa: UGATU, 2006. Vol. 8, № 1 (17). PP. 60-67.
5. Chomsky, N. “Language and a problem of knowledge”. The Bulletin of the Moscow State University: col. proceedings Moscow, 1996. Rel. 6. PP. 157-185.
6. Pogonin V.A. “Model of dispatching management robots”. Internet - magazine Information processes and management, 2006

Estimation of impact factors on regional informatization level

Albert Abilov
Izhevsk State Technical University
Izhevsk, Russia
e-mail: albert.abilov@mail.ru

Abstract

In this article the problem of digital divide has been defined and importance for impact estimation of social-economic factors on regional informatization development in different countries is justified. The impact of economic level (by GDP per capita) and education level (by Educational Index) on the two informatization indicators are explored: Internet users density and Broadband subscribers density. The Spearmen's rank method is employed for this research.

1. Introduction

Analysis of global dynamics reveals an acute problem of uneven informatization and economic resource distribution between the developed, developing and poorest countries. The problem has been called "a digital divide". The problem concerns the unequal opportunities for various sections of population, organizations and countries in getting access to advanced information and communication technologies. In many respects, the digital divide is caused by social factors and engineering capabilities. Both the living standards and the access of people to information technologies in the developed countries (such as the USA or West-European countries) have huge gap against the poor countries. Identification of crucial factors and estimation of their influence on informatization development both world-wide and at a country level becomes an actual problem under the conditions of digital divide. Basic factors influencing the informatization development are economics, education, technology and service. In addition, a lot of these factors intertwine in the field of informatizaton development. And so it becomes necessary to analyze their joint influence at quantitative level.

2. Impact Factors

Analysis of global dynamics reveals relationship between level of national economy and informatization development in the country. Table 1 shows averaged indicators of informatization and economic for three groups of countries: developed countries, developing countries and poorest countries at the beginning of 2009. A basic indicator of the level of economic development in a country is Gross Domestic Product per capita (GDP per capita). For the indicators of informatization level the follow values are used: Internet users density (IUD) and Broadband subscribers density (BSD).

Table 1 Informatization and economic indicators

Indicators	Developed countries	Developing countries	Poorest countries
IUD	68.0	40.8	4.5
BSD	24.8	9.6	0.1
GDP per capita	48.2	11.2	0.8

The data above indicate a stable dependence between quantitative informatization indicators and national economics. The higher is GDP per capita, the larger are informatization level. However, important as it is, economics is not the only factor influencing informatization. Potential users of information facilities can be grouped into four categories:

- users having need and access opportunities;
- users having need but no access opportunities;
- users having access opportunities but no need in them;
- users having neither access opportunities nor any need in them.

The first group is the best case. The presence of the second group is caused mainly by economical and engineering factors. Appearance of the third and the fourth groups results from the lack of motivation for information consumption. Economical factors are not the only cause of the problem, it is contributed by other social factors, such as level of educational attainment etc. One can conjecture that there exists a direct relation between demand for information service and educational attainment of society. High level of educational attainment and high qualification of citizens are key factors for growth of national economy and living standards. They are transformed to goods and services, contribute to efficiency of public sector, consolidate civil society and improve the investment climate. High-quality, equal and efficient education and science are important components of the transformation. In this connection it seems important to investigate standards of education in various countries,

since it is one of the key factors influencing informatization development.

One of the most famous indicator of education level is Education Index (EI). It is most often used in annual Reports on Human Development (UNDP) and some other analytical documents and reports. EI is a component of Human Development Index (HDI) and is calculated according to UNDP calculation procedure [1]. Education index is an inertial characteristic. Its behavior in many respects is determined by the policy of state, the amount of investments in education, the cultural aspects and many other factors. It should be also taken into account, that return of investments may be expected not earlier than in 10–15 years.

3. Economic and Informatization

Investigation of mutual influence of economics and info-communication is of great importance. Correlation functions are widely used for analysis of parameter interrelation. The functions use least-squares method and are based on Jipp diagrams [2]. Lines of correlation (Jipp diagrams) based on International Telecommunication Unit (ITU) data and connecting basic informatization indicators with national economy (GDP per capita) are shown in Fig. 1–2 (data of 2008, for 170 countries) [3]. The diagram points mark positions of various countries (some of them are labeled). The diagrams are double logarithmic plots.

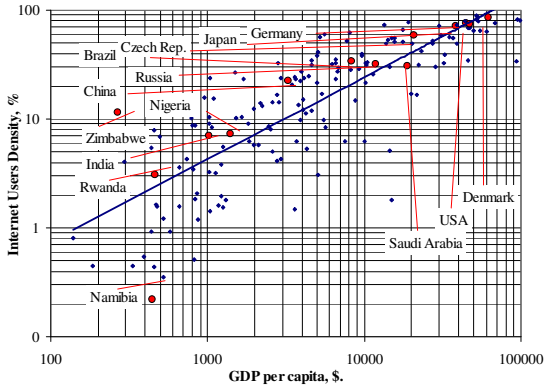


Figure 1. Interrelation of Economics and Internet users density

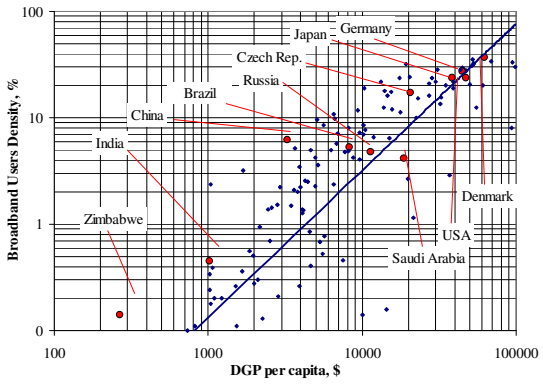


Figure 2. Interrelation of Economics and Broadband users density

Jipp diagrams demonstrate mutual influence of informatization and economics. The higher is economic level of the country, the higher is the average level of its informatization development. The relationship can be seen in the approximating curve. However, significant deviations from approximating curve are observed for some countries. Most likely, that may be attributed to the influence of some other factors. It is obvious, that by analysis of informatization development one cannot ignore the influence of external factors, such as economics, education, policy, population density etc.

For mathematical proof of interrelation availability between external factors and the level of info-communication development it is convenient to use Spearmen ranking method. The method enables to estimate, how close the parameter interrelation is. Spearman correlation coefficient (Spearmen Index) ρ is determined by the following expression [4]:

$$\rho = 1 - \frac{6 \cdot \sum_{k=1}^{n} (R_i - R_j)^2}{n(n^2 - 1)}, \qquad (1)$$

were k – sequence number of country; n – number of countries under examination; R_i, R_j – country ranks according to respective indicators.

The country rank is calculated for each parameter under examination according to the following principle: the higher is the parameter value, the higher is the country rank. Spearmen Index values lie in the range from 0 (no relation at all) to 1 (country ranks for both indicators coincide). The values are calculated on basis of statistical analysis for 170 countries all over the world. It is interesting to analyze the dynamics (development in time) of the influence of external factors on informatization level. In Fig. 3 values of Spearmen Index are shown. They enable to estimate relationship between informatization indicators (IUD and BUD) and economics in time.

As the services just appeared in the 90-s of the 20th century (initial stage), its Spearmen Index was low (in comparison with the present value): 0.5 and 0.7 for Internet users density (IUD) in 1990 year and Broadband users density (BUD) in 2000 year, respectively. So, Spearmen Index is low at the initial stage of new technologies and services formation. This implies presence of some other factors promoting creation and application of the technologies and services.

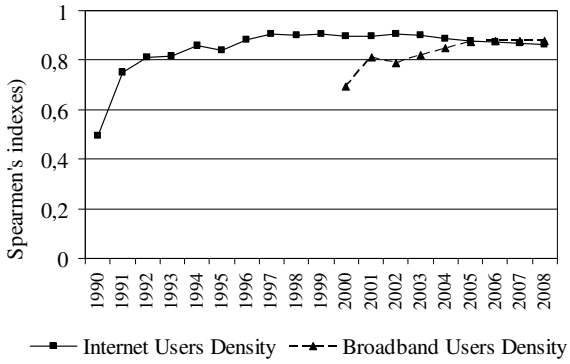


Figure 3. Dynamics of correlation between informatization and economic level (Spearmen Index)

Subsequent growth, followed by stabilization of the index value (over 0.9), indicates that economics is an important factor promoting development of information technologies and services. It is also important to examine behavior in time of education level influence, as it belongs to the factors influencing info-communication growth and development.

4. Education and Informatization

Advanced and available national education on the one hand and effective informatization development on the other hand can help the developing countries and the countries with transition economy to advance in stability and sustainable development of their socio-economic sphere. It is important to examine education level in various countries and regions and to analyze its influence on development of information technologies and services. Figs. 4 and 5 show correlation lines (Jipp diagrams) based on data of ITU and United Nations Development Programme (UNDP).

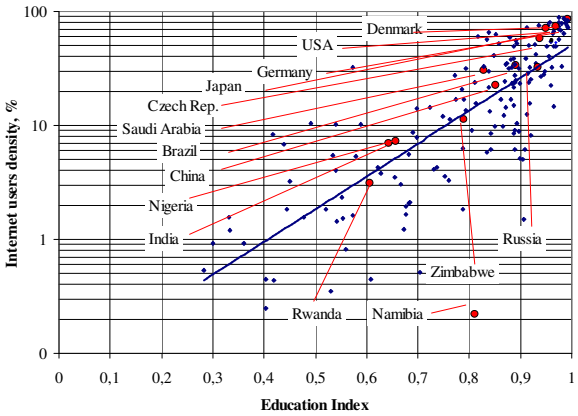


Figure 4. Interrelation of Education and Internet users density

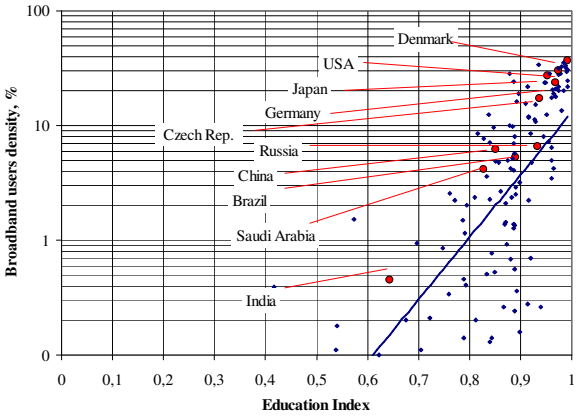


Figure 5. Interrelation of Education and Broadband users density

The correlation lines connect two informatization parameters (at the year-end of 2008) with UNDP Education Indexes (EIs) of 170 countries (at the begin of 2008) [1, 3]. For clarity of presentation the informatization axis has logarithmic scale, and EI axis – linear scale.

Statistical data show stable relationship between quantitative info-communication indicators (TLD, MCD, IUD и BSD) and Education Index of the country. Wide spread of points around approximating lines may be attributed to the fact that education and informatization cannot be considered separately from progress in other industrial sectors and from the level of economic development of the country as a whole. The constructed lines characterize only relationship with Education Index, whereas cases of wide point spread around the lines may be explained by presence of other influencing factors in respective countries. Saudi Arabia is an illustrative example: its Education Index is relatively low, so one should assume that high informatization indices of the country are, to a great extent, a consequence of high GDP per capita – 18855$ (in 2008). In 2007 the highest Education Index (0.993) was reported in five countries: Australia, Denmark, Finland, Ireland and New Zealand. The lowest EI (0,301) was in Byrkina Faso. Fig. 6 shows a bar chart of Spearmen Index values. The chart allows comparative analysis of interrelation strength between various types of information technologies and Education Index or GDP per capita.

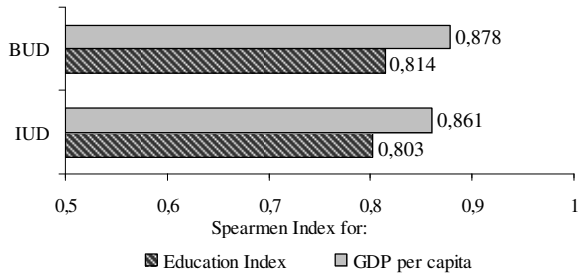


Figure 6. Spearmen index values

5. Conclusion

When considering the informatization development trends it is necessary to take into account the socio-economic impact factors. It is obvious, that informatization indicators depend on economic factors to a greater extent, but Education Index also exercise significant influence. Education Index and other education level indicators are quantitative characteristics. They ignore qualitative aspect which, undoubtedly, exercise influence on informatization and is determined by equipping level of schools and universities, professional skill of teaching staff, development of laboratories, extent of research work etc. High quality of education and high level of economics promote rapid and dynamic development of informatization in the country, facilitate its joining the global information society and are the main factors influencing informatization on national and global scale. This is why the factors may be considered as key factors influencing informatization.

References

17. Human Development Report 2009. Overcoming barriers: Human mobility and development. UNDP. 2009.
18. Jipp A. Welth of Nations and Telephone Density. Telecommunications Journal. 1963. – № 6.
19. Yearbook of Statistics – Telecommunication/ICT Indicators – 1999-2008. 35th Edition , 2009.
20. *Shmojlova R.A..,Shuvalova E.B., Glubokova N.Y.* Statistics Theory. – Moscow.: Finance and Statistics, 2001.

Integrated System for Controlling the Efficiency of Organizational Structures

Roman Fominykh, PhD Assoc. Prof.
STU Izhevsk (Votkinsk Branch)
Russia, Votkinsk
e-mail: saprlab@ya.ru

Abstract

The article is devoted to develop dynamic multifactor model of the functioning of industrial organizational structures. The model allows to evaluate and predict the effectiveness of ongoing organizational and technical measures. Evaluation is made on the analysis of the current state of organizational and technical level of production system, the level of real wages, constructive-technological complexity of manufactured products (services). Accounting for such endogenous indicators allows us to consider the organizational structure as a closed single unstructured whole, characterized by organizational and technical level, which produces a single universal product, measured in units of constructive-technological complexity, and the result of which is the level of real salaries of workers. The basis of the model laid a dynamic model of the economy with discrete time Solow, industrial-motivational concept G. Samarin, synergistic benchmarking - the practical application of scientific school "Economic Synergetics" B. Kuznetsova, the theory of constructive-technological complexity - the practical application of scientific direction "Theory of efficiency and improvement of integrated industrial production," the head of B. Yakimovich. The text of your abstract should be placed here. It is in your own interest to ensure the abstract adequately describes the content of your paper.

Introduction

After the Second World War in the U.S. have been 12 and 6 of the financial and economic crises. Total direct losses only to the U.S. economy, according to the group of scientists, G. Samarin, S. Doroshko, V. Chekirdy [1], amounted to $ 60 trillion dollars. And in view of lost profits - more than 100 trillion dollars. And this is only the United States. Given the loss of all developed countries, now humanity would long to solve the problems of poverty, environment and noosphere development.

Currently, all the leaders of developed countries have recognized that the existing economic models and theories do not allow them to effectively manage the economy. Practice shows that most economic theories can only give rise to crises.

Even more severe difficulties survive micro economy. C formidable challenges in planning its activities in the context of globalization, hypercompetition, the dominance of knowledge economy and the rule of innovative technologies facing most organizational structures (businesses, factories, enterprises, firms, etc.). Failure to capture the complexity of the problems, understand relationships and interactions, this is the main drawback of most managers. This also explains the actions of many large organizations and power structures that resemble the "blind flight".

Publishing

As a possible alternative paradigmal management organizational structures at the present time scientists STU Izhevsk (Votkinsk Branch) developed an integrated model for assessing and predicting the effectiveness of planned organizational and technical measures conducted within the organizational structure. The system aims to become one of the main tools to support decision making in the planning of organizational and technical measures to improve the functioning and competitiveness of organizations.

As the base of the model used in the dynamic model of economy with discrete time American economist Nobel laureate Robert Solow. Solow model describes the economy as a nonlinear dynamic system and is used to study transient processes in the economy associated with the replacement of one variant of macroeconomic policy to others.

In general, the developed model can be represented as follows:

$$
\begin{cases}
Y_t = F_3(T_t, Reg_t), \\
Y_t = I_t + S_t, \\
T_t = F_2(Q_t, U_t, C_t), \\
Reg_t = (1+\alpha) Reg_t \cdot k_3 \cdot k_5, \\
C_t = (1+\eta) C_{t-1} \cdot k_2, \\
Q_t = (1+q) Q_{t-1} \cdot k_1 \cdot k_4, \\
U_t = F_1(K_t, L_t), \\
L_t = (1+\nu) L_{t-1} + S_{t-1}, \\
K_t = (1+\mu) K_{t-1} + I_{t-1}, \\
k_1 = F_1'(K_{t-1}), \\
k_2 = F_2'(K_{t-1}), \\
k_3 = F_3'(K_{t-1}), \\
k_4 = F_4'(L_{t-1}), \\
k_5 = F_5'(L_{t-1}).
\end{cases} \qquad (1)
$$

where

t=0 — the base year;

t=T/ — The final year study period;

K_0, I_0, L_0, S_0, Reg_0, Q_0, C_0 — Are given;

Y — Financial performance (in monetary terms);

I — Volume of investment in fixed assets;

S — Compensation, including salary and all expenses of the company for the maintenance personnel, …;

K — Capital (fixed assets, stocks, debtors, cash, ….);

L — A generic indicator of personnel (in monetary terms);

U — Organizational and technical level of the organizational structure (support element);

T — Complexity of manufacture of the product in terms of organizational structure (support element), taking into account the organizational and technical level of the organizational structure, structural and technological complexity, and rate the quality of the project in its manufactured products;

Reg — Regional characteristics (income levels, minimum wages, unemployment, the consumer basket, the interest rate, inflation, level of education, monetary and fiscal policies and other socio-economic indicators of the region);

Q — An indicator of the quality of the project (design excellence);

C — Constructive-technological complexity of products, represents an inherent property of it, taking into account the geometric, structural and substantive properties of the product and its structural components, as well as placed upon them design and technological requirements in accordance with the existing level of development of productive forces.

k_1, k_2, k_3, k_4, k_5 — Coefficients considering the latent connection.

The basic idea of building an integrated model of motivation is the experience of Mr. Ford. H. Ford, as head of a company and not a professional economist, first in practice, found that pay is a major, major factor, in contrast to capital, technology, investment in determining economic growth, Gross domestic product, businesses of all sectors of the economy through their direct and latent communication and the entire socio-economic system [2]. Therefore, the presented model, attention is paid to the level true salaries S as a key competitive genesis [3].

Calculation of the index of organizational and technical level of the organizational structure U, complexity of manufacturing products in the organizational structure T and constructive-technological complexity of products C will be made on the basis of developed by scientists STU Izhevsk (Votkinsk Branch) methods [4,5,6].

Calculation of regional features Reg will be calculated based on the methodology proposed by G. Samarin, in developed by the industrial-motivational model [7].

Indicators characterizing the quality of the project (design excellence), determined in accordance with the methodology of synergistic benchmarking, the practical application of scientific school "Economic Synergetics" B. Kuznetsov [8].

The block diagram of a mathematical model for controlling the efficiency of organizational structures is shown in fig. 1.

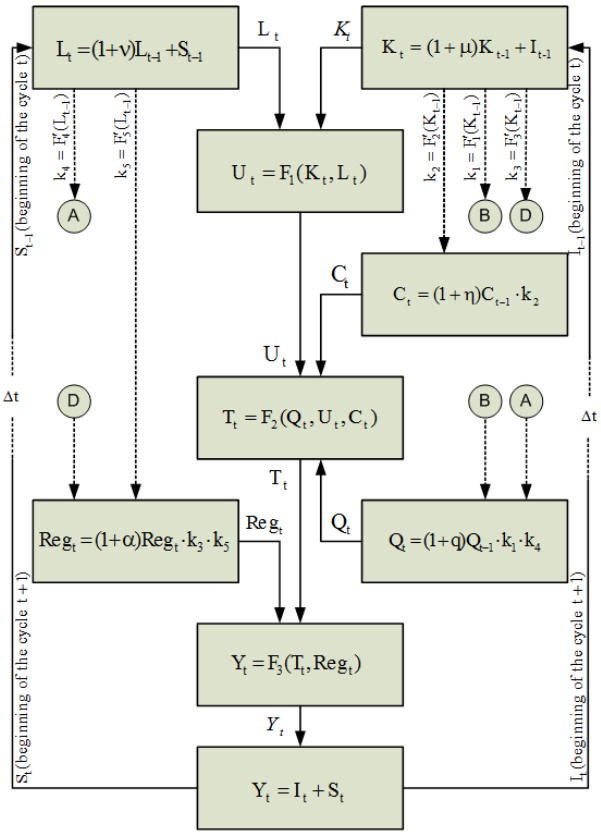


Figure 1. Block diagram of the mathematical model for controlling the efficiency of organizational structures.

Fig. 2 shows a graphical interpretation of the mathematical model for controlling the efficiency of organizational structures.

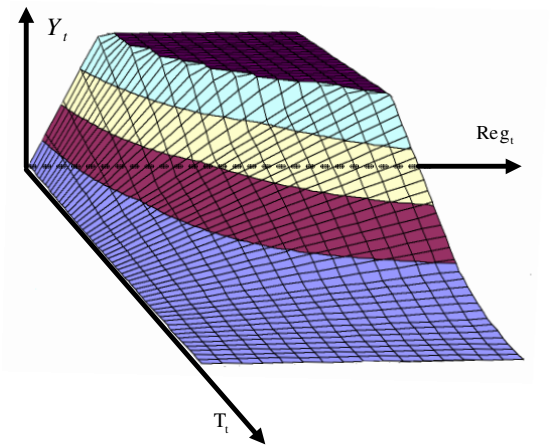


Figure 2. Graphical interpretation of the integrated system for controlling the efficiency of organizational structures

Conclusion

Designed on the basis of the mathematical model of an integrated system of performance management organizational structure will enable scientifically approach to strategic planning of organizational and technical measures aimed at ensuring the competitiveness of both the organization and the region as a whole.

References

1. Samarin, G.P., Doroshko S.E., Chekirda V.A. “Noospheric economy: back to basics. Basic value of work and motivation”. SPb.: PIF.com, 2008; pp. 338.
2. Henry Ford “My life, my accomplishments”. Ekaterinburg: Ural. JUDr. Publishing House: Vetrastar, 2003.
3. Fominykh R.L. “Methods of increasing the level of real salaries based on the theory of constructive-technological complexity and technological hype” In: *Proc. of the Eighth International Scientific and Practical Conference “Research, development and application of high technologies in the industry” High-tech, basic research, education and industry*, Spb.: Izd Polytechnic. Ca-ta, St. Petersburg, Russia, 2009, pp. 41-42.
4. Korshunov A.I., Fominykh R.L. “Using indicators of organizational and technical level of production system to assess the complexity of manufacturing production range” Intelligent systems in manufacturing: Periodic scientific-practical magazine, IzhSTU, Izhevsk, Russia, 2007, Vol. 1. pp. 128-138.
5. Fominykh R.L. “Development of an automated subsystems of constructive-technological complexity, complexity of manufacturing of parts and the organizational and technical level multinomenclature production” Thesis for the degree of candidate of technical sciences. IzhSTU, Izhevsk, Russia, 2003, pp. 134.
6. Korshunov A.I. “Creation of automated control systems of machinery production, based on the theory of constructive-technological complexity” Thesis for the degree of Doctor of Technical Sciences. IzhSTU, Izhevsk, Russia, 2008, pp. 313.
7. Galenko V.P., Samarin G.P., Strakhov O.A “Business planning in an open economy” A manual for university students, Publishing Center "Academy" Moskau, Россия, 2005.
8. Kuznetsov, B.L., Sharamko M.M. “Synergetic benchmarking” INEC, Neb. Chelny, Russia, 2006, pp. 187.

The improvement of the taxation of small-scale enterprises by means of information technology

Glukhova Maria
Blagodatsky Gregory
ISTU
Udmurt Republic, Russian Federation
e-mail: glukhovamn@yandex.ru, blagodatsky@gmail.com

Abstract[1]

Paper gives short review of corporate management system evolution from Grant's diagram to E-business [5, 6, 14, 18]. This paper also contains review of Russian Federation accounting information systems (Tester, Kamin, 1C Enterpise, Tax payer, Infin-managment). The advantages and disadvantages by using such systems were considered. Authors tries to give another view to fiscal effectiveness problem thought using new information system in cooperation with tax department.

1. Introduction

Most business leaders are trying to find a universal solution that can immediately eliminate the whole set of problems. Each year there is a new concept, the new method. However, there is one "but": there is no universal solutions. Everything depends on ourselves, and the real way is improvement of existing enterprise management approaches.

2. Paper text

The classical MRP II / ERP system involves relationship of planning and management functions with the functions of planning, accounting and orders, suppliers, production, customer, financial management. In the CIM system computer aided design systems (CAD systems) and process control system (PCS system) to all this are added. While earlier such a close interaction with the business system is not expected for these functions [15, 19]. Fig. 1 shows the general scheme of CIM.
Summary up, it is possible that CIM system integrates various software products, that have different ideology, operating systems and data formats. The very problem is the most difficult, but solvable. But, in practice, simple integration of different programs is not enough, because it must be cooperation and collaboration of enterprise divisions, that often conflict with each other.
To a large extent, the development of this technology was contributed by the establishment of international standards, which only recognizes some of the data exchange formats, and the creation of special programs - converters to facilitating information exchange. It is believed that in the future CIM will cover more applications developed in accordance with its standards or adapt to them. This process will continue so long as full and unrestricted interaction of any program from any manufacturers will be achieved.

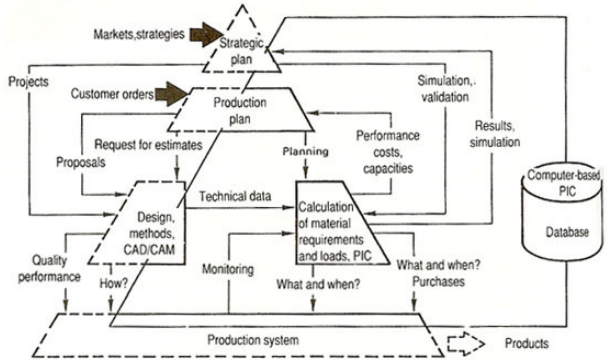


Figure 1. CIM general scheme

It is significant to say that the introduction of these systems are at a very slow. The reason is - the high cost of such decisions.
There are two management concepts in this area to date:
• Electronic business – ability to interacts with trading partners through Internet [14].
• CRM (Customer Relationship Management) is the management of interaction with customers [10].
These terms taken both individually and together sound quite attractive. The vendors of these systems say that they greatly increase the competitiveness of companies. However, they forget to say or deliberately silent that the possibility of quickly receiving orders through the Internet and information of the entire history of the relationship with the clients are not enough to deter them, or take a new one. There is a need for more reliable base, that including effective management system (JIT, MRP, MRP II, ERP, APS

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

and CIM). It allows to reduce the percentage of damage production, expedite fulfillment, improve quality, lower production costs. Using such system, the company can serve customers faster, while spending less resources and sell products at a lower price.
Finally most business leaders are trying to find a universal solution that can immediately eliminate the whole set of problems. Each year there is a new concept, the new method. However, there is one "but": there is no universal solutions. Everything depends on ourselves, and the real way is improvement of existing enterprise management approaches.

This may contribute to specific tools, managing information. These tools are:
MRP - the method of calculating necessary for the implementation of main production plan material's volume and component's quantity [2].
MRP II - the method of calculating required execution resources to perform an MPP (materials and components, production facilities and the ability of personnel, finance, etc.). The concept of management and class of enterprise information management systems that manage the main areas of company activity such as planning, production, finance, procurement and sales [8].
JIT - management philosophy, which is based on the principle: continually looking for processes that increase cost, while not increase the consumer cost of products, and eliminate them [3, 12, 16, 17].
TOC - the theory of constraints. There is the concept of management, administering "bottlenecks" [7, 8].
APS - a management method of new level. It allows quick intervening in the production process by making the necessary adjustments in the production plan of arbitrarily often. in time, it can displace MRP II [4, 11].
CRM - the management method, in account of the individual requirements of every customer to product. This technology allows to personalize products in detail [10].

ERP - enterprise management information system, that is based on MRP II management algorithms, implementing JIT and TOC methodologies. Modern ERP systems combination of APS and CRM modules and e-business solutions.
Announced the death of the above methods for over 10 years, but they are continue to be used. Indeed, today there aren't something new. We suppose that for Russian companies, their use is just in time now.
However when we speak about the tools, allowing to operate a volume of stocks, sales, a manufacturing process, we should not forget about a financial sphere supporting a business. Any small, average or large enterprise has to keep accounting records and to pay tax, a bookkeeper or an economist should use special information programs.
Let's begin with the program "Tax payer" - the program is developed by the Federal state enterprise "Main research computer centre of Federal Tax Department". The program allows to fill and send to tax departments electronic forms of the tax declarations. Programs "Tester" and "Kamin-tax" are similar programs, but "Kamin-tax" cannot work without "1C enterprise" platform. The program "Tester" is used for checking files conformity to formats of tax declarations and the accounting records in electronic form. Programs "Infin-management" and "1C Enterprise" are program complexes for managing various kinds of business. They allow to develop the unique decisions and they completely correspond to modern understanding ERP system. These programs are competitors having a similar production and working on one market. Large client an "Infin-management" is Lukoil. "1C Enterprise" has also large clients, for example, Beeline. Both platforms have their own models for tax payers in accordance with Federal Tax Department requirements. Programs "Tax payer", "Infin-management" and "1C Enterprise" allow to keep accounting records using traditional system of the taxation and special tax modes for a small-scale business.

These programs have both advantages and disadvantages. They help a bookkeeper to keep accounting records and fill tax declarations. But we should not to forget that quite often a self-employed businessman has no economic education and won't be able to understand all these tax rates, remissions, payment terms etc. Besides it is very expensive for self-employed businessman to apply even a reduced version of "1C Enterprise". For effective functioning small and average businesses it is necessary to further develop software regulating their activity and considering peculiarities of small business.
According to the law «On development of small and average business in the Russian Federation» the criteria of this business are:
- for small-scale enterprises an average number of workers is one hundred; for microenterprises - fifteen;
- for the microenterprises the maximum gain is 60 million rubles, for small enterprises – 400 million rubles.

It is necessary to develop a universal program for calculating taxes and filling tax declarations only for small business.
The remote system to control small business through the Internet in a real time from any point of the world and at any time should be developed in cooperation with tax department. The program should involve the basic functions of business — sales, purchases, the warehouse accounting, the finance, the personnel etc.
An enterprise needn't buy computer programs and an expensive server, lay and serve a network, pay services of IT experts and a bookkeeper. It will be enough to register an identification number, name and password of the tax payer on tax department site and the business control system will be ready to work.

By the results every day the businessman forms an extract about the incomes and expenses immediately through tax department base.
The information will be stored in a specialized data-centre with round-the-clock supervision and regular reservation. A computer and access to the Internet is everything that will be necessary. You can use any operating system — Microsoft Windows, Linux, MacOS; any web browser - Microsoft Internet Explorer, Mozilla Firefox, Safari, Opera, Google Chrome, Netscape.

3. Conclusion

On the whole, the system will have a number of obvious advantages for a small business:
- business managing at any time and from any place;
- instant receiving the up-to-date information on the taxation;
- refusing the purchasing of an expensive server and programs;
- decreasing expenses on IT service and the bookkeeper;
- automatic installing new versions of system;
- reliable and safe data storage.

Such a system is profitable to the state too, because tax department will be able to constantly monitor the development of small enterprises in regions and immediately make decisions the way to stimulate small business.

References

1. Advanced Planning and Scheduling Solutions in Process Industry / Edit. Hans-Otto Ganther, Paul van Beek. – Cambridge: Springer, 2003. – 426 p.
2. APICS dictionary / Edit. J. F. Cox, etc. – Chicago: American Production and Inventory Control Society, 1992. – 144 p.
3. Ford, Henry. My life and work. / Henry Ford. – Sioux Falls: NuVision Publications, 2007. – 192 p.
4. Gaither, Norman. Production and operations management / Norman Gaither, Gregory V. Frazier. – Cincinnati: South-Western College Publishing, 1999. – 846 p.
5. Gantt, Henry L. A graphical daily balance in manufacture. / Henry L. Gantt. // Transactions of the American Society of Mechanical Engineers. – 1903. – Volume XXIV, – P. 1322-1336.
6. Gantt, Henry L. Work, Wages, and Profits / Henry L. Gantt. – New York: Engineering Magazine Co., 1916. – 330 p.
7. Goldratt, Eliyahu M. The Goal: A Process of Ongoing Improvement / Eliyahu M. Goldratt, Jef Cox. – Great Barrington: North River Press, 2004. – 384 p.
8. Goldratt, Eliyahu M. The Race / Eliyahu M. Goldratt, Robert E. Fox. – Aldershot: Gower Publishing Ltd., 1994. – 181 p.
9. Keller, Erik L. Enterprise Resource Planning : The changing application model / Erik L. Keller // SAP R/3 3.1. – Stamrord: Gartner Group, 1996. – P.8.
10. Kincaid, Judith W. Customer Relationship Management : Getting It Right / Judith W. Kincaid. – New York: HP Books and Prentice Hall PTR, 2002. – 512 p.
11. Korem, Yoram. Computer Control of Manufacturing Systems / Yoram Korem. – New York: McGraw Hill Inc., 1983, – 304 p.
12. Liker, Jeffrey. The Toyota Way : 14 Management Principles From The World's Greatest Manufacturer / Jeffrey Liker. – New York: McGraw-Hill Inc., 2003. – 350 p.
13. Orlicky, J. Material Requirements Planning : The New Way of Life in Production and Inventory Management / J. Orlicky. – New York: McGraw-Hill Inc., 1975. – 292 p.
14. Papazoglou, Mike. e-Business : Organizational And Technical Foundations / Mike Papazoglou, Pieter Ribbers. – Indianapolis: John Wiley & Sons, 2006. – 750 p.
15. Rehg, James A. Computer-Integrated Manufacturing / James A. Rehg , Henry W. Kraebber. – New Jersey: Prentice Hall, 2004. – 592 p.
16. Taiichi, Ohno. Just-In-Time for Today and Tomorrow / Ohno Taiichi. – New York: Productivity Press, 1988. – 145 p.
17. Taiichi, Ohno. Toyota Production System : Beyond Large-Scale Production / Ohno Taiichi. – New York: Productivity Press, 1988. – 163 p.
18. Taylor, Frederick Winslow. The Principles of Scientific Management / Frederick Winslow Taylor. – New York: Dover Publications Inc., 1998. – 76 p.
19. Waldner, J. CIM : Principles of Computer Integrated Manufacturing / J. Waldner. – Indianapolis: J. Willey & Sons, 1992. – 206 p.

Customer-Orientation Using Integration and Individualization Aspects Enabling the Transition from Manufacturer to Solution Provider

Dipl. Wi.-Ing. Alexander Burger
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: alexander.burger@kit.edu

Dipl. Inform. Vitalis Bittel
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: vitalis.bittel@kit.edu

Dipl. Wi.-Ing. Ramez Awad
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: ramez.awad@kit.edu

M.Sc. Stilian Stanev
Institute for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: stilian.stanev@kit.edu

Prof. Dr. Dr.-Ing. Jivka Ovtcharova
for Information Management in Engineering
Karlsruhe Institute of Technology
Karlsruhe, Germany
e-mail: jivka.ovtcharova@kit.edu

Abstract[1]

The paper addresses the increasing importance of solution providers due to fundamental changes within global economic constraints. The challenges emerging during the transition process in becoming a solution provider will be highlighted as well as the lack of scientific approaches dealing with these challenges. Based upon the integration and individualization aspects of solutions, fields of action for future approaches will be depicted. In this context, the importance of continuous and individual customer integration will be accented, followed by a methodology facilitating a customer-individual value provision as well as the alignment of solution portfolios according to changing customer needs. The paper will conclude with the benefits of the methodology as well as with the emphasis of future need for action regarding customer integration in the field of solution provision.

1. Introduction

For some time past, industrial enterprises from different industries are confronted with a fundamental change of global economic constraints. This change is due to an aggravated international competition especially from low-wage countries and results in a homogenization of the offered products regarding technical and qualitative aspects [1, 2]. In consequence of the missing possibility to differentiate from competitors by means of the available product range, the market players engage in an intense price competition, which affects the providers' profit margins negatively. Empirical studies show that industrial enterprises understand the increasing pricing pressure due to a missing possibility of differentiation from competitors as a main problem area, which has to be faced necessarily [3].

Customers, especially within the business consumer sector, recognize the opportunity to increase their market power since providers of homogeneous products can easily be interchanged without the risk of establishing dependences. In doing so, the customers not only enforce decreasing market prices but also make growing demands concerning the services of the providers [4]. Those growing demands become manifest in the fact that customers claim offerings which are tailored to their very special needs [5]. The origin of those claims is in an increasing need of individualization and is thus embedded into a global development as individualization is considered as one of the main actual global trends.

Many manufacturers meet those challenges by extending their product-driven core business with services going along with the products. The aim of this strategy is to set up key differentiators from the

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

competition as well as to generate an individual added value for the customer by additional services. This step is accompanied by a changing self-conception of the companies - from a manufacturer to a solution provider. Hereby, a solution provider is defined as a provider of customized and customer-specific problem solutions whereas manufacturers are only and solely providers of produced goods [6]. Besides the opportunities to set up differentiation potentials and therefore increase the enterprise's market power as well as to meet customer requirements in an individual way, the entrance into the solution business allows an increase of the enterprise's profit margin. Thus, the attainable EBIT-margin of solutions comes up to 11.0% and therefore exceeds the margin of business with new machines, which is 2.3% exemplarily regarding plant construction, by far [7, 8].

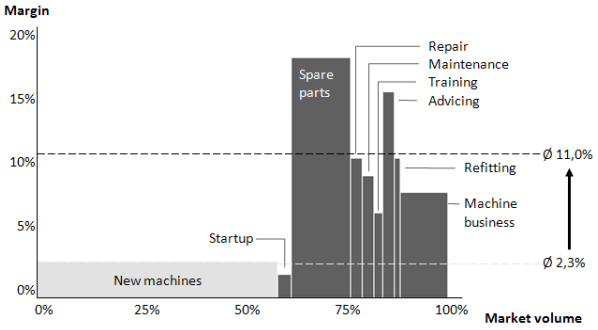


Figure 9 EBIT-margins and market volumes of solution components [8]

Empirical studies show that industrial enterprises understand the installation of such a solution business as a main challenge of the future [9]. Providers such as IBM, General Electric, or Rolls-Royce have already made the step towards a solution provider and successfully offer comprehensive customer solutions instead of isolated products and services [10].

2- Deficiencies in the transformation process becoming a solution provider

Since the entrance to the solution business happens in an intuitional and hardly structured way in many enterprises, the expected potentials are hardly exploited or even not at all. Therefore, the anticipated differentiation from competition fails to succeed in many cases [11, 12]. In certain cases, this development even leads to the quitting of the solution business. The reasons arise from a missing understanding of the idea of a comprehensive solution, but also from a lack of scientifical methods, which accompany the transition to a solution provider. In the following, the main reasons for the failure occurring during the transition to a solution provider will be summarized.

2.1 Inadequate methods supporting the development of services

Regarding manufacturers, the planning and development of the product business as a historical core competence is based on methodically established foundations and follows a structured way. In contrast, deficits can be found regarding the range of services since methods and tools for a structured planning and development of services are only provided to an inadequate degree. Therefore, the development of services in the industry happens in an unstructured way and ad hoc in response to specific customer requests. In doing so, the fact that the economic success of a service is significantly dependent on its conception and design is considered only in an inadequate way. Hence, in analogy to the development of products, services are to be regarded as development objects which allow a continuous and lifecycle-supporting handling and management based on information-technological reference models. The academic field of Service Engineering is build upon the mentioned basic ideas but is limited to the domain of services without taking the integration between products and services into consideration. Due to the short history of the discipline Service Engineering, in comparison to Product Engineering there are no standardized procedure models which can be applied by enterprises. Furthermore, the service sector is characterized by a lack of scientific approaches dealing with reference models [13].

2.2 Deficiencies within the integration of product and service domain

In the referred context of customer solutions, the limitation of focusing only to the domain of services as covered within Service Engineering is not sufficient. Instead, the lifecycles of the product and service domains have to be integrated as well as aligned into each other in order to meet the concerns of the solution aspect in a sufficient way. However, scientific research has considered theoretical insights and practical experiences about how enterprises can effectively apply their development resources so as to develop integrated solutions only for a short time and therefore only exist to a minor degree. Exemplarily, the approaches introduced by Zhang, et al. [14] as well as from Spath and Demuß [15] shall be named. However, up to now the focus of research is on the economic marketing of hybrid products, on business models as well as pricing theory and only to a lesser extent on the provision of consistent approaches for integrated development processes. According to Becker et al. [16], only three reference models dealing with integrated products and services exist, whereas only specific topics such as facility management are considered in these approaches.

In consequence of the lack of adequate methods, product and service processes in enterprises are mostly interconnected only to a minor degree. The resulting product-service-systems are characterized by a low degree of harmonization and contribute to the satisfaction of customer-specific requirements as well as to the differentiation from other market participants

only in an insufficient way. Simple but often offered services such as installation or maintenance are only partly sufficient within the present context. Many of the offered services only aim at the subsequent disposal of products and therefore the customers expect those services to be as a matter of course and for free. Instead, services which are appropriate to the problem and which individually target and permanently optimize the value-added processes of customers have to be developed and finally offered. Such services are characterized by their difficult substitutability and therefore allow an escape from the easy replaceability of the product range when the services are intimately involved into the product business. In contrast, the entrepreneurial reality appears as follows:

- The single parts of the solution are only cursorily integrated with each other and therefore easily substitutable.
- The customer solutions do not offer an added value compared to the separate purchase of product and service components.

2.3 Deficiencies in degree of individualization of the solution offering

The degree of individualization is a critical factor when deciding to what extent customer-specific but still enough standardized solution portfolios can be offered. The problem follows from the fact that in spite of the important role, which the provision of a customer-individual solution statement plays, an approach, which is effective and efficient for the service provider, is demanded. Due to these two restrictions, providers find themselves in the area of conflict between complete standardization and individual customization.

Standardized solutions are easily reproducible and can be sold in large amounts on one hand. However, they do not match the understanding of customer solutions since the individual customer requirements are not considered. Due to their simple reproducibility, the anticipated differentiation from the market is not possible. On the contrary, a complete customized individualization decreases the efficiency on the side of the service provider and contributes to the fact that the provisioning costs exceed the expected benefits many times over. Whereas the benefits for the customer as well as the possibility of differentiation amplify, profit margins and efficiency as well as reproducibility decrease due to higher provisioning costs.

Enterprises face this problem by providing a pre-defined solution portfolio with the possibility of customized specification. Thereby, the customer-specific adjustment is gained by configuration or mass customization, as the whole portfolio is subdivided into individual, partially freely selectable sub-components. However, the main challenge how to define the right granularity of the individual components still remains. Hereby, the definition of larger solution components tends to standardization since smaller but more individual components mean a continuously high maintenance effort as the range of solutions has to be adjusted to changing market requirements continuously.

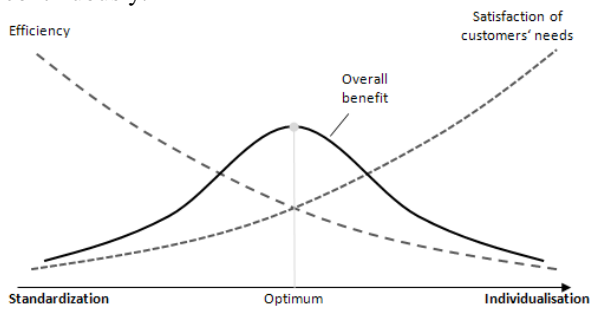


Figure 10 - Vendors in the area of conflict between standardization and individualization

2.4 Insufficient integration of individual customer requirements

Due to a lack of methodical support, solution providers do not integrate customers into the processes of solution creation and solution provision in a sufficient way, affected in particular are the design and development phases.

In the classical sense, the customer was faded out inside the processes of the service provider and was only considered as a passive recipient [17]. For solution providers with the deep-seated aspiration to solve customer-individual problems, the integration of single customers becomes compulsory inevitable. The necessity results in the heterogeneity of customer-individual problems which can only be solved by means of customer-specific solution alternatives.

Already existing methods from Product Engineering as well as from Service Engineering provide only insufficient support since the customer only partially needs to be integrated into the production process, especially regarding Product Engineering. Thus, Berkovich in her state-of-the-art analysis also emphasizes the necessity to integrate customers into the whole development process and points out that no existing approach of customer integration covers the whole innovation process [18]. Usually, the customer integration on the level of Product Engineering is limited to the general requirements management as well as to the sales process. In the following, existing methods of customer integration are outlined.

In the quality management standard DIN EN ISO 9000:2005, a requirement is defined as „a need or expectation which is determined, usually preconditioned, or compulsory." Within the context of product development, the requirements management enables the customers and also internal development departments to communicate their expectations on products which are to be developed in the future. For this purpose, Product Engineering offers several process models such as the iterative approach of Ehrlenspiel [19] or the sequential process model of Pahl [20]. The available approaches integrate the

prospective customer into the production processes only by explicit questionnaires or checklists but do not show possibilities how requirements can be passed to the provider in an intuitive manner using the natural process integration of the customer.
Within Service Engineering, methods, processes, and tools of the requirements management are only covered superficially. In the process model of Ramaswamy, the requirements of some exclusive key users are understood as crucial. However, it is not mentioned how the process of requirement collection should happen [21]. Schneider et al. also mention the importance of customer requirements but do not offer appropriate methods [22].
In contrast to the requirements management which defines expectations on new or more customized products and services, Feedback Management applies itself to the return of experiential knowledge about existing services to the development processes. Feedback from customers can be converted into requirements subsequently to its acquisition. Referring to Ovtcharova [23], Feedback Management is defined as methodology which aggregates all specialized activities of an enterprise concerning the return of experiential knowledge into the development processes of products and services which become more adjusted to customer requirements. The process contains the acquisition, aggregation as well as the integration of the conditioned knowledge into the operational processes of product and service planning and development.
Both presented methods have in common that they do not consider the continuous integration of customers but instead focus on selected aspects concerning customer integration. The allocation of the specific requirements of individual customers is only of secondary interest. Instead, an anonymous aggregation of customer requirements is the basis of prospective development processes. Thus, a tailored provision of customized solution alternatives is not possible with any of the available methods due to the lack of a customer-specific problem analysis as well as the lack of customer-specific requirements concerning products or services.

3. Objectives and associated work plan

In the preliminary chapters it was illustrated that companies changing from sole manufacturers to solution providers are confronted with manifold challenges, which shall be analyzed systematically in the following. As a basis for the analysis, the definition of the term "solution" will be referred, in order to better demonstrate the causes for the uprising problems and subsequently identify adequate possibilities for the avoidance of those problems based on methodical approaches.
The relevant literature distinguishes solutions from simple product-service-bundles using three central characteristics. According to the derived definition, a solution is (1) a combination of products and services, which (2) may be adapted to special customers' wishes (3) and are closely linked with each other [11, 24, 25, 26]. By means of the presented definition, the two main aspects of the term solution can be clarified. These main aspects are integration and individualization.

3.1 Integration aspect of solutions

Concerning solution offers, the relevant literature distinguishes between two fundamental kinds of integration: marketing-based integration on one hand and technical integration on the other hand. The marketing-based integration enables customers to obtain the solution, which is appropriate to their problem, directly from a single source. Thus, instead of multiple vendors, the solution provider emerges as a general contractor, whereby consumers obtain an added value throughout the entire decision and acquisition cycle [25]. In this context, literature mentions cooperative arrangements of solution providers with suppliers as well as competitors, enabling them to offer a wide range of services most suitable for the customers' specific problem statement [10, 27].
Technical integration on the other hand describes to what extent product and service components within a solution offering are inter-coordinated and harmonized. Hereby, it is not sufficient to offer services which are largely detached from products. Instead, the development of the individual components of the product and service domain has to be carried out in such a way, that an added value for the customers arises. Thus, as objective of the integration of single components, the value of a solution perceived by the customer has to exceed the sum of its single elements. [28]
In combination, both kinds of integration illustrated in the previous sections ensure that a solution provider is able to differentiate from other market participants by the provision of more integrated and thus less substitutable solution offerings. The general rule is, that the higher the degree of integration in technical and marketing-based kinds are pronounced, the higher the potential to differentiate compared to the competitive environment is recognized [29].
The general objective arising from the integration aspect of the solution term is to offer customers a solution value, which exceeds the sum of the values of the single solution components, by the combination of product and service components. For this purpose, the product and service domain have to be integrated into each other. Beyond, it has to be assured that all solutions or solution components required for the fulfillment of the customer's necessities are offered from one single source. The knowledge, which solutions have to be provided for the fulfillment of the problem-specific necessities, may be acquired by the individualization aspect of solutions (see chapter 3.2).

3.2 Individualization aspect of solutions

The individualization aspect of solutions addresses the ability of solution providers to satisfy customer-specific necessities on the basis of a pre-defined solution portfolio as comprehensive as possible. Thereby, the integration of individual customers is not limited to the integration of the customer within the scope of the solution specification. Instead, the customer as an equal partner within the value chain has to be integrated at an early stage and continuously involved into the processes of the solution provider. Therefore, customers have to be integrated into all stages of an ideal process of solution provision, which are namely analysis, solution specification, solution provision and solution result [30].

- Analysis of the customer's necessities: On the part of the provider there is the need of understanding each customer's specific problems and necessities. For many customers, the mutual analysis of the specific customer problem already belongs to the problem-solving, for what reason the customer's early and continuous integration is once again emphasized to be mandatory.
- Solution specification: Subsequently, a problem-adequate definition of the solution based on the mutual analysis of the requirements becomes essential. The definition likewise takes place with the customer's interaction and has to be carried out in a transparent and comprehensible way. The provision of a solution portfolio adequate to the customer's needs has to be assured.
- Solution provision and solution result: In a manufacturer's product business, the lifecycle for the producer mostly ends with sales activities. In contrast, solutions are characterized by the fact that the provision of the solution is also carried out by the solution provider. Thereby, the solution provider becomes directly responsible for the incidence of the desired impact, namely the solution of the customer's problem. The solution provision is accompanied by a constant alignment with the solution result and therefore also by the involvement of the customer. This approach facilitates the solution provider to counteract by adjusting the solution specification if necessary.

As objective arising from the individualization aspect of the solution term, the demand for an individually customizable, pre-defined solution portfolio arises. This solution portfolio has to be aligned to the satisfaction of customer-individual needs while taking into account the shifting nature of the customers' challenges. The provision of such a solution portfolio is based on methods and tools for the continuous integration of individual customers into the processes of the solution provider.

4. Solution approach

Resulting from the objectives of the solution's integration aspect, the demand for an integrated planning, development, and provision of solution arises. Hence, processes of product and service domain have to be harmonized pursuing the main goal to provide a coherent solution portfolio which is value-adding for the customer. A reference model which includes the product domain as well as the service domain serves as a basis for the integration and therefore links both domains with each other. Such a reference model has to be developed as a component-based, hierarchically structured model with the possibility of customer-individual specification which allows the development and provision of a pre-defined solution portfolio. Exemplarily, such a model was developed and adopted in collaboration with a manufacturing company within the scope of an industrial cooperation. The reference model for services originated in this context is subdivided into the three partial models of the result dimension, the process dimension and the potential dimension analogous to the three dimensions of constructivist service definitions. The integration into the product domain occurs via the process modules, which are interlinked with assemblies and maintenance positions. **Figure 11** exemplarily illustrates the partial model of the result dimension.

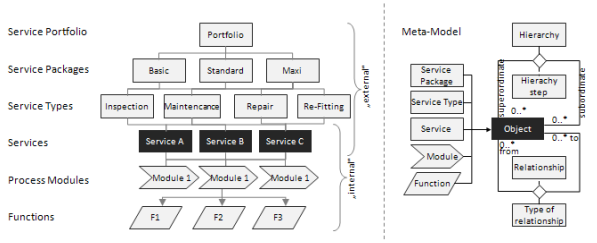


Figure 11 - Result dimension of the service reference model

The individual turnout of solution alternatives requested in the final objectives of the individualization aspect is solved on the basis of configuration management using mass customization. Following the approach of the 'product instance PLM' described by Abramovici et al., [31] a concept for the customer-specific instantiation of the general reference model is suggested, building the foundation for the customer integration method described in the following. Dependent on customer specific projects, instances of the general model are created allowing the deposition of relevant field information. The relationship between instances and the general model are persistent in order to ensure the possibility of bi-directional information exchange.

A three-level methodology for customer integration permits the continuous integration of individual customers into the processes of the solution provider based on customer requirements as well as on customer feedback. The methodology assumes that customers continuously emit implicit and explicit information referring to the solution definition of the provider. Based on this information, the solution provider is enabled to (1) determine the exact customer-specific

solution configuration on one hand and on the other hand to (2) continuously align the solution portfolio to the necessities of the customers.

- Within the first level of the methodology, the information emitted by the customer is collected and connected with the customer instance of the general reference model. The assignment to individual components of the instantiated model offers the possibility of a solution configuration characterized by transparency and adjustment to the customers' needs by the analysis of the customer's requirements interconnected with the instantiated model.
- In a second step, the aggregation of the customer-specific information and the subsequent assignment to certain components of the general reference model in the context of requirements management allow the adjustment of existing solution portfolios to shifting market needs. Therefore, customers' requirements on a solution configuration as well as requirements extracted by practical knowledge (customer feedback) lead to the enhancement of existing solutions.
- Beyond, the aggregation of customer information without the possibility of assignment to existing solution components allows the planning of new product and service components in the context of innovation management. Frequently mentioned, but due to the lack of corresponding components not grantable customer requirements thereby are included into the planning processes of new solution components.

Such a customer-centric approach both increases the innovation capabilities of an enterprise as well as the customer retention, as always those solutions can be offered which assure the highest potential benefit for the customer. **Figure 12** illustrates the methodology explained above.

5. Conclusion and outlook

Within this paper it was shown that the change from manufacturer to solution provider entails manifold challenges. Based upon the definition of the solution term, operational fields and solution alternatives supporting companies in their transformation process were identified. Particularly, the necessity of integration between product and service domain was highlighted and a reference model for the provision of pre-defined solution portfolios was introduced. Beyond that, the paper emphasized the importance of continuous customer integration and pointed out the lack of existing approaches. It was stated that the deficiencies in this field especially are affecting innovation processes of solution providers. As a solution approach covering this problem, a three-layer methodology for customer integration was illustrated, facilitating a customer-individual value provision as

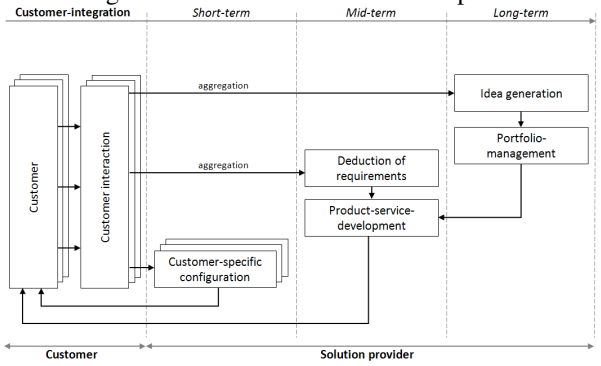


Figure 12 - Proposed methodology for continuous customer integration

well as the alignment of solution portfolios according to changing customer needs. Future research will have to pursue integrating the customer into the processes of the solution providers enabling them to serve clients in a more individual and customer-oriented way. Therefore, standardized integrated reference models are needed which are, above all, dealing with individualized customer information. Models enriched with semantic information show great promise for this problem as the interconnection between the different domains becomes obvious and interdependencies can be depicted [32].

Acknowledgements

The authors wish to acknowledge their gratitude and appreciation to all research partners for their valuable cooperation and their contribution, especially to Prof. J. Ovtcharova for her contribution during the development of various ideas and concepts presented in this paper as well as to Ms. V. Altmann for her support.

References

1. Homburg, C.; Faßnacht, M.; Günther, C. Erfolgreiche Umsetzung dienstleistungsorientierter Strategien von Industriegüterunternehmen. Zeitschrift für betriebswirtschaftliche Forschung. 2002, Bd. 54, 6
2. Günther, C. Das Management industrieller Dienstleistungen: Determinanten, Gestaltung und Erfolgsauswirkungen. Wiesbaden : Gabler, 2001
3. McKinsey Corporation. The McKinsey Global Survey of Business Executives. The McKinsey Quarterly. 2005, Bd. 42, 1
4. Stremersch, S.; Weiss, A.; Dellaert, B.; Frambach, R. Buying Modular Systems in Technology-Intensive Markets. Journal of Marketing Research. 2003, Bd. 40, 3
5. Reichwald, R.; Bonnemeier, S. Kommunikation in der Wertschöpfung. [Buchverf.] M. Bruhn, F.-R. Esch und T. Langner. Handbuch Kommunikation. Grundlagen, innovative Ansätze, praktische Umsetzungen. Wiesbaden : Gabler, 2009
6. Schuh, G.; Friedli, Th.; Gebauer, Th. Fit for Service: Industrie als Dienstleister. München, Wien : Hanser, 2004

7. Gebauer, H.; Hildenbrand, K.; Fleisch, E. Servicestrategien für die Industrie. Harvard Business Manager. 2006, Bd. 27, 5
8. Roland Berger Strategy Consultants. Maschinen- und Anlagenbau im Abschwung. September 2008
9. Sturm, F.; Bading, A. Investitionsgüterhersteller als Anbieter industrieller Lösungen – Bestandsaufnahme des Wandels anhand einer Umfrage. Wirtschaftsinformatik. 2008, Bd. 50, 3,
10. Davies, A.; Brady, T.; Hobday, M. Charting an Path Toward Integrated Solutions. In: MIT Sloan Management Review. 2006, Bd. 47, Nr. 3, S.39-48
11. Johansson, J.; Krishnamurthy, C.; Schlissberg, H. Solving the Solutions Problem. The McKinsey Quarterly. 2003, Bd. 40, Nr. 3, S. 116-125
12. Neu, W.A.; Brown, S.W.: Forming Successful Business-to-Business Services in Goods-Dominant Firms. In: Journal of Service Research. 2005, Vol. 8, No. 1, S. 3-17
13. Fähnrich, K.-P., van Husen, C.: Entwicklung IT-basierter Dienstleistungen: Co-Design von Software und Services mit ServCASE, Physica-Verlag, 2008
14. Zhang, Y.; Srai, J; Gregory, M.; Iakovaki, A. Engineering Network Configuration: Transition from products to services. In: Proceedings of the 1st CIRP Industrial Product-Service Systems (IPS2) Conference, Carnfield, Carnfield University Press, 2009
15. Spath, D.; Demuß, L. Entwicklung hybrider Produkte – Gestaltung materieller und immaterieller Leistungsbündel. Bullinger, H.-J.;Scheer, A.-W. Service Engineering – Entwicklung und Gestaltung innovativer Dienstleistungen. 2006S.463-502
16. Becker, J.; Beverungen, D.; Knackstedt, R. Reference Models and Modeling Languages for Product-Service-Systems Status Quo and Perspectives for Further Research. In: HICSS '08: Proceedings of the 41th Annual Hawaii International Conference on Systems Sciences, Washington. IEEE Computer Society. 2008, S.105-115
17. Prahalad, Coimbatore K.; Ramaswamy, Venkat: The Future of Competition: Co-Creating Unique Value with Customers. Boston, MA: Harvard Business School Press. 2004, S.2
18. Berkovich, M.; Esch, S.; Leimeister, J.M.; Krcmar, H. Requirements engineering for hybrid products as bundles of hardware, software and service elements – a literature review. 9. Internationale Tagung der Wirtschaftsinformatik. Wien, Österreich, 2009
19. Ramaswamy, R. Design and Management of Service Processes, Addison-Wesley, Massachusetts, 1996
20. Ehrlenspiel, K. Integrierte Produktentwicklung, Hanser, München, 2003
21. Pahl, G.; Beitz, W.; Feldhusen, J.; Grote, K.H. Konstruktionslehre, Berlin: Springer, 2003.
22. Schmitz, G. Die Ermittlung der Kundenanforderungen an industrielle Dienstleistungen, in: Zeitschrift für Planung. 2000, Nr. 2, S. 195-215
23. Ovtcharova, J.; Krahtov, K.; Rogalski, S. eHomeostasis Methodology in the Automotive Industry - Integration of Emotion-based Customer Feedback into Product Lifecycle Management, in: Proceedings of the MCPC 2007 World Conference on Mass Customization & Personalization (MCP), Innovation and Research Conference, MIT Cambridge/Boston, 2007
24. Davies, A.; Brady, T. Organisational Capabilities and Learning in Complex Product Systems: Towards Repeatable Solutions. Research Policy. 2000, Bd. 29, 7-8
25. Sawhney, M. Going Beyond the Product: Defining, Designing and Delivering Customer Solutions. [Buchverf.] R. Lusch und Stephen L. Vargo. Toward a Service-Dominant Logic of Marketing: Dialogue, Debate, and Directions. Armonk, New York, 2006, S. 365-380
26. Tuli, K.; Kohli, A.; Bharadwaj, S. Rethinking Customer Solutions: From Product Bundles to Relational Processes. Journal of Marketing. 2007, Bd. 71, 3
27. Windahl, C.; Andersson, P.; Berggren, C.; Nehler, C. Manufacturing Firms and Integrated Solutions: Characteristics and Implications,in: European Journal of Innovation Management. 2004, 7 (3), S. 218-228
28. Böhmann, T.; Krcmar, H. Hybride Produkte: Merkmale und Herausforderungen. [Buchverf.] M. Bruhn und B. Stauss. Wertschöpfungsprozesse bei Dienstleistungen. Wiesbaden : Gabler Verlag, 2007
29. Krishnamurthy, C.; Johansson, J.; Schlissberg, H. Solutions Selling. Is the Pain Worth the Gain? McKinsey Marketing Solutions. McKinsey Marketing & Sales Practice. 2003, S. 3
30. Scheer, A.-W.; Grieble, O.; Klein, R. Modellbasiertes Dienstleistungsmanagement in: Bullinger, H.-J.; Scheer, A.-W. Service Engineering – Entwicklung und Gestaltung innovativer Dienstleistungen 2006, S.22-24
31. Abramovici, M.; Fathi, M.; Holland, A.; Neubach, M. Integration von Feedbackdaten aus der Produktnutzungsphase im Rahmen des PLM-Konzepts In: Proceedings zur Multikonferenz Wirtschaftsinformatik (MKWI 2008), München, Germany, 2008
32. Bittel, V.; Awad, R.; Ovtcharova, J. Analysis and semantic structuring of product information with the example – distribution, Proceedings of the Russian-German Workshop "Innovation in Information Technologies". Ufa, Russia, July 25-31, 2009

Modelling of Organizational Structure of the Enterprise: the Process Approach

R.A.Badamshin
Ufa state aviation technical University I
Ufa, Russia

L.N.Rodionova
Ufa state aviation technical University I
Ufa, Russia
e-mail: rodion@ufanet.ru

T.A. Sultanov
Ufa state aviation technical University I
Ufa, Russia

Abstract[1]

Approaches to formation and to an estimation of multilevel structures of management of the enterprises are examined in the article , the analysis of existing methods of an estimation of influence of risk factors on realization of projects on perfection of control systems is carried out. Methodical recommendations according to expediency of introduction of the process approach.

Any enterprise constantly solves a two-in-one problem - moves to achievement of the strategic targets, and adapts to environment change. Thus development long-term objectives remain invariable, and ways of their achievement can vary according to change of macro - and microenvironment. Construction of a control system by changes consists in creation of multilevel structure of management with use of the process approach.

The analysis of ways of representation of various, poorly formalizable data [1, 2] shows, that for the solution of a problem of formation of private and integrated indicators of functioning MSU and its elements it is possible to take advantage of substantive provisions of the theory of serving system (SS) and theories of mark off processes which allow to describe the stochastic phenomena and processes effectively.

The mathematical analysis of SS work is very simplified, if process of this work is Markov's. For that end it is enough, that all streams of events translating system from a condition to a condition (streams of demands, «service streams»), were elementary. The casual process running in the system, is called *Markov's* if for any moment of time t_0 probabilistic characteristic of process in the future depend only on its condition at present t_0 and do not depend on, when and how the system came to this condition.

Let at the moment t_0 the system is in certain condition S_0. We observe process otth (on the third hand) and at the moment t_0 we know a condition of system S_0 and all background of process, everything, that was at $t < t_0$. The future of ($t > t_0$) interests us .It is impossible to foresee it exactly, as the process is accidental, and it means that it is unpredictable. But we can find probabilistic characteristic of the process in the future. For example, the probability that after some time t the system S turns out to be S_1 or will keep position S_0, etc.

In practice Markov's processes in the pure state usually are not found, but we have to not quite often to deal with processes for which "background" influence can be neglected. Studying such processes it is possible to apply Markov's models successfully.

The important characteristic of a flow of events is its *intensity* - the average of events falling to unit of time. Intensity of a stream can be both as constant, so variable, depending from time t.

For the simplest stream with intensity λ the interval between the next events has so-called exp distribution with density

$$f(t) = \lambda e^{-\lambda t}, (t > 0). \quad (1)$$

Size λ in the formula (1) is called as parameter of the exponential rule. For random value T having exp distribution, the population mean m_T is a quantity, back to parameter, and average quadratic deviation σ_T is equal to the population mean:

$$m_T = \sigma_T = 1/\lambda \quad (2)$$

In the theory of probability as «accident measure» non-negative incidental is quite often examined so-called factor of a variation:

$$v_T = \sigma_T / m_T \quad (3)$$

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

From formulas (2), (3) follows, that for exponential distribution the coefficient of variation is to 1, that is for the elementary stream of events the factor of variation of intervals between events is equal to unity.
It is obvious, that for a regular stream of events at which an interval between events is not casual, the variation factor is equal to zero.
Let consider system S having n of possible conditions S_1, S_2 ,..., S_n. We name probability of i-th condition probability $p_i(t)$ that during the moment t the system will be in condition S_i. It is obvious, that for any moment the sum of all probabilities of conditions is equal to unity.
Having at one's disposal the marked state graph, it is possible to find all probabilities of conditions $p_i(t)$ as time functions. So-called Kolmogorov's formulas - the differential formulas of a special kind in which unknown functions are state probability for this purpose are made and solved. Each member is equal to product of density of probability of transition of the corresponding arrow multiplied on probability value of those of conditions from which the arrow proceeds.
Realization of the project of decision-making on conversion of new production in structure of operation of business can be schematically presented each, as consecutive connection multichannel of SS from unlimited long turn and unlimited expectation (fig. 1).

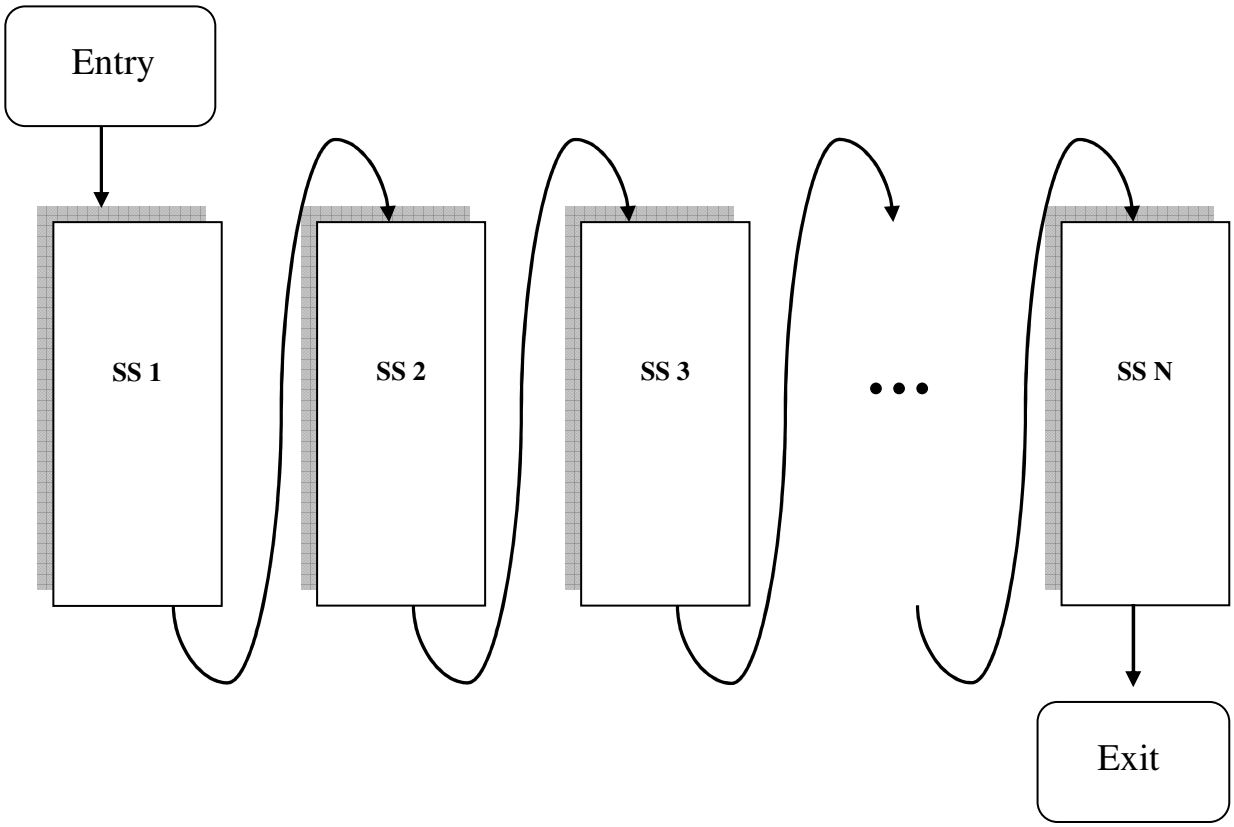


Figure 1. The General Scheme of Consideration of the Project in Functional Structure of Management

Let's define the basic characteristics of the represented system consecutive SS. We will examine n-channeled SS on which the stream of demands with intensity λ (i.e. the average of the demands arriving in unit of time, is equal λ) arrives; intensity of service for one channel μ (μ - size, return to an average holding time $t_{обсл}$); the number of places in turn is unlimited. The demand which has found all channels occupied, stands in line, where there are no imposed restrictions neither on length of turn, nor on a waiting time.
The states of a system are numbered according to number of the demands connected by system:
There is no turn: S_0 - all channels are free; S_1 - one channel is occupied, all the others are free; … S_n - all channels are occupied; there is a turn; S_{n+1} - all channels are occupied,one demand stands in line ; … S_{n+m} - all channels are occupied, m demands stand in turn.
Graphic representation of such SS is presented in figure 2. On arrows from left to right the system is translated always by the same stream of demands with intensity λ, on arrows from right to left the system is translated by the service stream which intensity is equal μ, increased by number of the occupied channels.

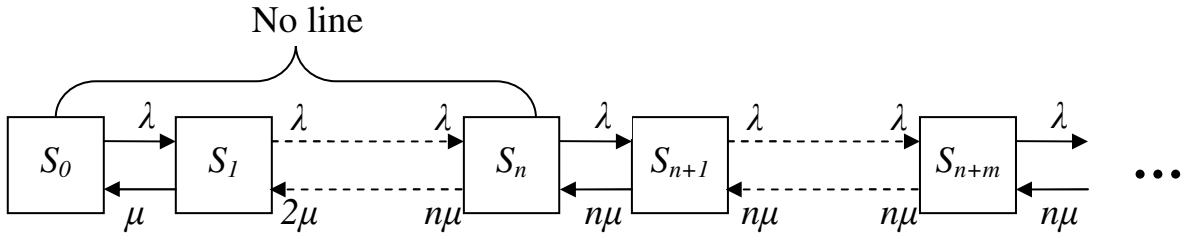


Figure 2. n-Channeled System of Mass Service with Unlimited Long Turn and Unlimited Expectation

As basic characteristics of SS we will examine:

$P_{отк}$ - probability of that to the demand will be given up in service;

$P_{обсл}$ - probability of that the demand will be served (or relative capacity), $P_{обсл} = 1 - P_{отк}$;

A - absolute capacity, i.e. an average of the demands served in unit of time, $A = \lambda \cdot P_{обсл}$;

$\bar{z}$ – an average of the occupied channels, $\bar{z} = \frac{A}{\mu} = \frac{\lambda}{\mu} \cdot P_{обсл}$;

q – a share of the channels occupied with service, $q = \frac{\bar{z}}{n}$;

$\bar{r}$ - an average of demands in turn;

$\bar{k}$ - an average of demands in system, $\bar{k} = \bar{r} + \bar{z}$;

$\bar{t}_{ож}$ - an average waiting time of the demand in turn, $\bar{t}_{ож} = \frac{\bar{r}}{\lambda}$;

$\bar{t}_{обсл}$ - an average holding time of the demand, $\bar{t}_{обсл} = \frac{P_{обсл}}{\mu}$;

$\bar{t}_{СМО}$ - Average time of stay of the demand in SS, $\bar{t}_{СМО} = \bar{t}_{ож} + \bar{t}_{обсл}$.

For reception of characteristics of the process proceeding in SS we will examine at first *n*-channel CMO from the limited long turn (fig. 3).

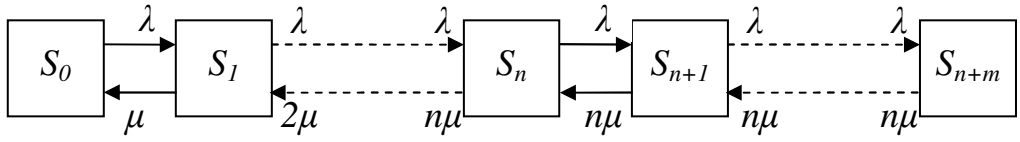


Figure 3. *n*-Channeled SS from the Limited Long Turn

Kolmogorov's formula for probabilities of conditions examined of SS will become:

$$\begin{cases} \frac{dp_0(t)}{dt} = \mu\, p_1(t) - \lambda\, p_0(t), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dp_k(t)}{dt} = \lambda\, p_{k-1}(t) + (k+1)\mu\, p_{k+1}(t) - (\lambda + k\mu) p_k(t), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dp_n(t)}{dt} = \lambda\, p_{n-1}(t) - n\mu\, p_n(t), \\ \frac{dp_{n+1}(t)}{dt} = \lambda\, p_n(t) - n\mu\, p_{n+1}(t), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dp_{n+m}(t)}{dt} = \lambda\, p_{n+m-1}(t) - n\mu\, p_{n+m}(t). \end{cases} \quad (4)$$

The equations (4) are called as Erlang's formulas. Natural entry conditions for their decision are:

$$p_0(0) = 1,\ p_1(0) = p_2(0) = \ldots = p_n(0) = p_{n+1}(0) = \ldots = p_{n+m}(0) = 0$$

which correspond to that during the initial moment the system is free.

Integration of system of the formula (4) in an analytical kind is difficult enough; therefore in practice such systems of the differential equations usually dare numerically. The probability bound of the conditions characterizing installed operating mode of SS have the greatest interest for practice. For finding of probability bound we will take advantage of already solved problem received for the scheme of destruction and reproduction [3,4,5,6,7]:

$$\begin{cases} p_k = \dfrac{\rho^k}{k!} p_0, \quad k = \overline{1,n}; \\ p_{n+i} = \dfrac{\rho^{n+i}}{n^i n!} p_0, \quad i = \overline{1,m}; \\ p_0 = \dfrac{1}{1 + \dfrac{\rho}{1!} + \dfrac{\rho^2}{2!} + \ldots + \dfrac{\rho^n}{n!} + \dfrac{\rho^{n+1}}{n\,n!} + \dfrac{\rho^{n+2}}{n^2\,n!} + \ldots + \dfrac{\rho^{n+m}}{n^m\,n!}} = \\ = \dfrac{1}{1 + \dfrac{\rho}{1!} + \dfrac{\rho^2}{2!} + \ldots + \dfrac{\rho^n}{n!} + \dfrac{\rho^n}{n!} \cdot \dfrac{\dfrac{\rho}{n} - \left(\dfrac{\rho}{n}\right)^{m+1}}{1 - \dfrac{\rho}{n}}} \end{cases} \quad (5)$$

Intensity of a stream of demands λ and intensity of a stream of services(for one channel) μ do not appear in these formulas separately, and enter only relation λ/μ.

We will designate this relation: $\rho = \lambda/\mu$

Also we will name size ρ «the resulted intensity» stream of demands. Its physical sense is that: size ρ represents an average of the demands coming in SS for an average holding time of one demand.

Let's define system effectiveness characteristics.

Probability of refusal. The arrived demand is refused, if all *n* channels and all *m* places in turn are occupied :

$$P_{отк} = p_{n+m} = \frac{\rho^{n+m}}{n^m\,n!} p_0$$

.

Relative throughput supplements probability of refusal to unit:

$$P_{обсл} = 1 - P_{отк} = 1 - \frac{\rho^{n+m}}{n^m\,n!} p_0$$

.

Absolute throughput SS:

$$A = \lambda P_{обсл} = \lambda\left(1 - \frac{\rho^{n+m}}{n^m\,n!} p_0\right)$$

.

Average of the occupied channels:

$$\bar{z} = \frac{A}{\mu} = \frac{\lambda}{\mu}\left(1 - \frac{\rho^{n+m}}{n^m\,n!} p_0\right) = \rho\left(1 - \frac{\rho^{n+m}}{n^m\,n!} p_0\right).$$

The average of demands in turn can be calculated directly as a population mean of a discrete random variable:

$$\bar{r} = 1 p_{n+1} + 2 p_{n+2} + \ldots + m p_{n+m} = \sum_{i=1}^{m} i \cdot \frac{\rho^{n+i}}{n^i \cdot n!} p_0 =$$

$$= \frac{\rho^{n+1}}{n \cdot n!} p_0 \sum_{i=1}^{m} i \cdot \left(\frac{\rho}{n}\right)^{i-1} = \frac{\rho^{n+1}}{n \cdot n!} p_0 \sum_{i=1}^{m} i \cdot \chi^{i-1}$$

Where $\chi = \rho/n$.

For definition of sum $\sum_{i=1}^{m} i \cdot \chi^{i-1}$ we will take advantage of a derivative of the sum of a geometrical progression:

$$\sum_{i=1}^{m} i \cdot \chi^{i-1} = \left(\sum_{i=1}^{m} \chi^i\right)' = \left(\frac{\chi\left(1 - \chi^m\right)}{(1 - \chi)}\right)' = \frac{1 - \chi^m\,(m + 1 - m\chi)}{(1 - \chi)^2}$$

Thus:

$$\bar{r} = \frac{\rho^{n+1} p_0}{n\,n!} \frac{1 - \chi^m\,(m + 1 - m\chi)}{(1 - \chi)^2}$$

Average of demands in system: $\bar{k} = \bar{r} + \bar{z}$.

Let's define an average waiting time of the demand in turn. We will examine a number of the situations, differing themes, in what condition again come demand will find system and what is the time it should wait for service.

If the demand finds not all channels occupied, it should not wait (corresponding members in a population mean are equal to zero). If the demand comes during the moment when all *n* channels are occupied, and the turn is not present, it should wait on the average time, equal to $1/n\mu$ (because «a stream of clearings» *n* channels has intensity $n\mu$). If the demand finds all channels occupied and one demand before itself in turn, it should wait on the average during time $2/n\mu$ (on $1/n\mu$ standing demand everyone ahead) etc. If the demand finds in turn k of demands, it should wait on the average during time $k/n\mu$. If again come demand finds in turn already *m* demands she will not wait at all (but also it will not be served). An average waiting time we will find, multiplying each of these values on corresponding probabilities:

$$\bar{t}_{ож} = \frac{1}{n\mu} p_n + \frac{2}{n\mu} p_{n+1} + \ldots + \frac{m}{n\mu} p_{n+m-1} =$$

$$= \frac{1}{n\mu}\left(\frac{\rho^n}{n!} p_0 + 2\frac{\rho^{n+1}}{n n!} p_0 + \ldots + m\frac{\rho^{n+m-1}}{n^{m-1} n!} p_0\right) =$$

$$= \frac{\rho^n p_0}{n\mu n!} \sum_{i=1}^{m} i \cdot \left(\frac{\rho}{n}\right)^{i-1} = \frac{\rho^n p_0}{n\mu n!} \sum_{i=1}^{m} i \cdot \chi^{i-1} = \frac{\rho^n p_0}{n\mu n!} \frac{1 - \chi^m(m + 1 - m\chi)}{(1 - \chi)^2}$$

Average time of stay of the demand in system: $\bar{t}_{СМО} = \bar{t}_{ож} + \bar{t}_{обсл}$.

For calculation of characteristics of SS from unlimited long turn and unlimited expectation we will use received before parity at $m \to \infty$.

Probabilities of conditions we will receive from formulas (5) limiting transition (at $m \to \infty$). Thus we will notice, that the sum of a corresponding geometrical progression converges at $\chi < 1$ and disperses at $\chi > 1$. Therefore requirement $\chi = \frac{\rho}{n} < 1$ is obligatory for reception of final values of probabilities of conditions. Thus, expressions for limiting probabilities of conditions will become:

$$\begin{cases} p_k = \dfrac{\rho^k}{k!} p_0, \quad k = \overline{1,n}; \\ p_{n+i} = \dfrac{\rho^{n+i}}{n^i n!} p_0, \quad i = 1,2,\dots; \\ p_0 = \dfrac{1}{1 + \dfrac{\rho}{1!} + \dfrac{\rho^2}{2!} + \dots + \dfrac{\rho^n}{n!} + \dfrac{\rho^{n+1}}{n!(n-\rho)}} \end{cases}$$

In SS from unlimited long turn and an unlimited waiting time each demand will be served sooner or later, therefore $P_{обсл} = 1$, and $P_{отк} = 0$, $A = \lambda \cdot P_{обсл} = \lambda$, on the average continuously occupied channel will give out $\lambda/\mu = \rho$ the served demands in unit of time.

The average of demands in turn can be calculated directly as a population mean of a discrete random variable:

$$\bar{r} = \sum_{i=1}^{\infty} i \cdot p_{n+i} = \sum_{i=1}^{\infty} i \cdot \frac{\rho^{n+i}}{n^2 \cdot n!} \cdot p_0 = \frac{\rho^{n+1}}{n \cdot n!} \cdot p_0 \cdot \sum_{i=1}^{\infty} i \cdot \left(\frac{\rho}{n}\right)^{i-1} =$$

$$= \frac{\rho^{n+1}}{n \cdot n!} \cdot p_0 \cdot \sum_{i=1}^{\infty} i \cdot \chi^{i-1}.$$

Expression $\sum_{i=1}^{\infty} i \cdot \chi^{i-1}$ represents a derivative of the sum of an infinite geometrical progression. Considering requirement $\chi = \rho/n < 1$, the sum of infinitely decreasing geometrical progression is equal $\sum_{i=1}^{\infty} \chi^i = \frac{\chi}{1-\chi}$, and its derivative looks like $\frac{1}{(1-\chi)^2}$. Thus, $\bar{r} = \frac{\rho^{n+1}}{n \cdot n!} \cdot p_0 \cdot \frac{1}{(1-\chi)^2}$.

Using similar reasoning, we will receive an average waiting time: $\bar{t}_{ож} = \frac{\rho^n p_0}{n \mu n!} \frac{1}{(1-\chi)^2}$. The Average of the occupied channels: $\bar{z} = \frac{A}{\mu} = \frac{\lambda}{\mu} = \rho$. The Average of the demands connected with SS: $\bar{k} = \bar{r} + \bar{z}$.

On the basis of resulted above formulas for calculation of basic characteristics of SS with unlimited turn and an unlimited waiting time all characteristics of the process presented on fig. 3 can be calculated. Every private SS possesses the parameters and characteristics. However, considering, that on the average continuously occupied in every SS the channel will give out $\lambda/\mu = \rho$ the served demands in unit of time, all SS will give out $\rho \cdot n$ demands in unit of time which will arrive in following SS. Thus, intensity of receipt of demands in each subsequent of SS will depend on two sizes: intensity of receipt of demands in the first SS and from amount of given out demands in unit of time previous SS, namely, intensity of receipt of demands in each the subsequent SS is defined at least from the specified two sizes. Really, if previous SS processes demands faster, than they arrive, average amount of demands in unit of time, arriving in the subsequent SS, cannot be more than amount of the demands arriving in unit of time in system. The example of calculation of characteristics of process is presented in table 1.

Table 1. The Example of Calculation of Characteristics of Process

Parameters of SS	SS 1	SS 2	SS 3	SS 4	SS 5	SS 6	SS 7
n	4	4	5	4	5	5	5
λ	0,1	0,1	0,1	0,1	0,1	0,1	0,1
μ	0,09	0,11	0,11	0,13	0,25	0,20	0,20
ρ	1,10	0,90	0,90	0,80	0,40	0,50	0,50
χ	0,28	0,23	0,18	0,20	0,08	0,10	0,10
$P_{отк}$	0	0	0	0	0	0	0
$P_{обсл}$	1	1	1	1	1	1	1
A	0,1	0,1	0,1	0,1	0,1	0,1	0,1
$\bar{z}$	1,1	0,9	0,9	0,8	0,4	0,5	0,5
q	0,275	0,225	0,18	0,2	0,08	0,1	0,1
$\bar{r}$	0,0106	0,0042	0,0005	0,0024	0,0000	0,0000	0,0000
$\bar{k}$	1,1106	0,9042	0,9005	0,8024	0,4000	0,5000	0,5000
$\bar{t}_{ож}$	0,1060	0,0416	0,0054	0,0240	0,0001	0,0002	0,0002
$\bar{t}_{обсл}$	11	9	9	8	4	5	5
$\bar{t}_{СМО}$	11,106	9,042	9,005	8,024	4,000	5,000	5,000

Summing up all elements of lines "$\bar{t}_{ож}$", "$\bar{t}_{обсл}$" and "$\bar{t}_{СМО}$", it is possible to calculate the time of waiting, service and stay of the demand in all SS. For the examined example the listed sizes have made 0,177, 51 and 51,177 time units accordingly. Besides it, the calculated characteristics of all SS give the information which is a basis for acceptance of administrative decisions.

Thus, it is possible to assert, that application of methods of the theory of Markov's processes and systems of mass service allows to solve effectively problems of perfection of organizational structure of operation of business of communication with the risk account.

Acknowledgement

The research is supported by grant of President of Russian Federation № 65497.2010.9 for Leading Scientific Schools.

References

1. Mike Robson, Phillip Ullah. Practical guidance on business process reengineering. - M: Audit, «Unity», 1997. - 224 p.
2. Pol H.Allen. Bank reengineering: the survival and success program / translated from english - M.:Alpina Publisher, 2002. - 264 p.
3. Gnedenko B.V., Kovalenko I.N.The introduction in the theory of mass service. The 4 publish. - M: Publishing house LKI, 2007. 400 p.
4. Kleinork L.The theory of mass service. - M: Mechanical engineering, 1979. - 432 p.
5. Bocharov P. P, Pechinkin A.V.The theory of mass service. - M: RUDN, 1995. - 530 p.
6. Matveev V. F, Ushakov V.G.Sistemy of mass service. M: Moscow State University Publishing house, 1984. - 240p.
7. Kofman A.The methods and models of research of operations. M: Mir, 1966. - 524.

Information Technologies Contribution in Client-Oriented Marketing

A. Hilbert
Fakultät Wirtschaftswissenschaften
Technische universität Dresden
Dresden, Germany
e-mail: hilbert@wiid.wiwi.tu-dresden.de

D.R. Bogdanova
Department of computer science and robotics
Ufa state aviation technical university
Ufa, Russia
e-mail: dianochka7bog@mail.ru

N.V. Sholokhova
Department of computer science and robotics
Ufa state aviation technical university
Ufa, Russia
e-mail: nadya_88@inbox.ru

Abstract[1]

The need to take into account individual preferences and technological limitations predetermines its multi-criterion nature. The completion of such task would be impossible without the application of intelligence information technologies. In this paper described an approach that takes into account uncertain information on clients and their individual preferences. It is hoped that this discussion assists marketers in forging a solid base for understanding and executing customer segmentation.

1. Introduction

Client-oriented marketing is the part of marketing directed on your "best customers" - those who make your basic profit.

As well as in many aspects of marketing, there is no «correct way» or «a wrong way» for marketing: what is necessary to do will depend on the purposes you wish to reach. The purposes are always the same: to increase quantity of "the best customers" and profit, brought by them. It is simple idea, but till now many companies lose huge profits without giving a proper attention to the constant clients.

One of the methods of client-oriented approach is to create the constant program of loyalty:

The given method assumes the use of the program of loyalty for creation of base of constant clients, their segmentation into groups, and then use of different marketing strategies for each segment. This approach allows you to learn your clients in more details, to use target marketing, and you will have an information how to modify your business according to the desires of customers. [1]

Performing marketing research, it often appears a problem of processing the great volume of the qualitative (i.e. subjective, uncertain) information on clients. To solve such problems can be used methods of fuzzy logic which refer to methodology of an artificial intelligence.

In this paper will be described two of these methods.

RFMR method

RFMR is a scoring approach which assesses recency, frequency and monetary ratio of a customer's purchases along an interval scale and aggregates the results to a single customer value score. This score can be used to define customer segments and develop marketing plans.

10. Recency is the time that has elapsed since the customer made his most recent purchase. A customer who made his most recent purchase last month will receive a higher recency score than a customer who made his most recent purchase three years ago.

11. Frequency is the total number of purchases that a customer has made within a designated period of time. A customer who made six purchases in the last three years would receive a higher frequency score than a customer who made one purchase in the last three years.

12. Monetary Ratio is each customer's average purchase amount. A customer who averages a $100 purchase amount would receive a higher monetary ratio score than a customer who averages a $20 purchase amount (average purchase amount = total dollars spent on purchases in last three years / total number of purchases in last three years) [5].

The purpose of RFMR scoring is to drive better segmentation decisions by valuing customers along an

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

interval scale. The most common approach is to sort customers in descending order (best to worst). Customers are broken into five equal groups (quintiles). The best receive a score of 5, the worst a score of 1. For recency, customers are sorted by days since last purchase, the lower number of days – the higher the score. For frequency, customers are sorted by number of purchases, the higher number of purchases the higher the score. For monetary ratio, customers are sorted by the amount spent. The higher amount the higher the score. Each time customers are scored, a new relative segmentation scheme is created. This has the advantage of quantifying customer behaviour which can be projected into the future. The relatively best customers would always fall into the 5, 5, 5 category. It is necessary to identify where the cut-off points fall, because they automatically change with every new customer. This kind of sorting can be applied if one deals with quantitative variables [3]. However, in case of fuzzy sets related to linguistic variables, sorting by membership function value should be applied.

For example:

1. Customer 1: Recency (R) - $\mu_{1high} = 0.8$, $\mu_{1low} = 0.3$;
   Frequency (F) - $\mu_{2high} = 0.6$, $\mu_{2low} = 0.5$;
   Monetary R. (MR) - $\mu_{3high} = 0.7$, $\mu_{3low} = 0.4$.
2. Customer 2: Recency (R) - $\mu_{1high} = 0.4$, $\mu_{1low} = 0.6$;
   Frequency (F) - $\mu_{2high} = 0.5$, $\mu_{2low} = 0.4$;
   Monetary R. (MR) - $\mu_{3high} = 0.8$, $\mu_{3low} = 0.3$.
3. Customer 3: Recency (R) - $\mu_{1high} = 0.9$, $\mu_{1low} = 0.2$;
   Frequency (F) - $\mu 2_{high} = 0.3$, $\mu_{2low} = 0.4$;
   Monetary R. (MR) - $\mu_{3high} = 0.1$, $\mu_{3low} = 0.6$.

Then, by means of the function max (μhigh, μlow), the highest membership function value is selected.

Customer 1: Recency (R) – $\mu_{1high} = 0.8$;
Frequency (F) - $\mu_{2high} = 0.6$;
Monetary R. (MR) - $\mu_{3high} = 0.7$.

Customer 2: Recency (R) – $\mu_{1low} = 0.6$;
Frequency (F) - $\mu_{2high} = 0.5$;
Monetary R. (MR) - $\mu 3_{high} = 0.8$.

Customer 3: Recency (R) – $\mu_{1high} = 0.9$;
Frequency (F) - $\mu_{2low} = 0.7$;
Monetary R. (MR) - $\mu_{3low} = 0.9$.

Now, all customers can be sorted by each indicator separately (R in ascending order, F in descending order, MR in descending order). The following sequence will result:

Recency (R): Customer 2 ($\mu_{1high} = 0.6$);
Customer 1 ($\mu_{1high} = 0.8$);
Customer 3 ($\mu_{1high} = 0.9$).

Frequency (F): Customer 1 ($\mu_{2high} = 0.6$);
Customer 2 ($\mu_{2high} = 0.5$);
Customer 3 ($\mu_{2low} = 0.4$).

Monetary R. (MR): Customer 2 ($\mu_{3high} = 0.8$);
Customer 1 ($\mu_{3high} = 0.7$);
Customer 3 ($\mu_{3low} = 0.6$).

As an example, two parameter values of the indicators (high and low) were used, and actually there could be more parameter values (very high, high, middle, low, very low etc.). The customer quintile method has the advantage of yielding equal numbers of customers in each segment. There are five equally-sized groups for recency, frequency and monetary ratio (according to a principle of Pareto, 20 % of customers bring 80 % of profit, hence, quintiles are used), generating 125 segments overall (cells would have definitions like: 4, 3, 5 or 2, 3, 3) [3]. Borders of each segment are calculated by dividing the ordered list of the customers into the desirable number of quantiles (n, usually five). Then, a score from 1 to n is assigned to the different segments. As an example, 3 segments are defined:

Recency (R): Customer 2 - 3 points;
Customer 1 - 2 points;
Customer 3 - 1 point.

Frequency (F): Customer 1 - 3 points;
Customer 2 - 2 points;
Customer 3 - 1 point.

Monetary R. (MR): Customer 2 - 3 points;
Customer 1 - 2 points;
Customer 3 - 1 point.

Since past behaviour can be regarded as the best predictor of future behaviour, recency is typically considered as the most powerful of the three variables. Many direct marketing decisions are solely based on recency. Unlike frequency and monetary ratio, customers reset themselves. A three-year-long reordering customer who has purchased an average amount only once, for example, moves up from 1 to 2 regarding frequency and monetary ratio, whereas regarding recency, he may move from 1 to 5. Customers that order regularly may hardly have anything other than a 5 score. At the core of recency is the idea that most of the customers fall into two groups: hot and dead. With relational, database-driven marketing databases becoming more common, most marketers can select RFMR scores independently. However, aggregating the scores allows for an overall segmentation of the customers based on a combination of different characteristics. When working with linguistic variables,

one alternative for the aggregation would be simply to add together the RFMR scores discussed above. The best customers would have a composite score of 15 (5+5+5) and the worst customers would have a minimum score of 3 (1+1+1). Many of the customers would have a score of 7 or 8 and it would be difficult to put them into an order. Moreover, the experience of direct mail marketing suggests that the most recent customers are of greater value than those who have ignored more than a few repeated mailings. To enhance the aggregation, the scores are often multiplied with different weights, e. g. Rx3, Fx2 and MRx1. This would give the best customers a composite score of 30 (5x3 + 5x2 + 5x1). This not only draws more attention to the most recent activities, it also gives a bit of a boost to frequency. The idea behind weighting frequency is that if two customers have equal recency, spent the same amount but one ordered several times and the other only once, the more frequent buyer is much more likely to respond. One additional enhancement is often employed by creating a composite score using the weighting factors 9.9, 6.6 and 3.3 instead of 3, 2 and 1. This yields a range of composite scores between 99 and 19.8 and preserves the approximately 3x weighting of R, while it also creates more of a 100 point scale.

Operational management of services based on individual preferences of customers

In an ever increasing competition and customers' raised requirements there arises a problem of enhancing operational management efficiency of services provided, which is a multifactor task.

The need to take into account individual preferences and technological limitations predetermines its multi-criterion nature. The completion of such a task would be impossible without the application of intelligence information technologies.

The paper looks at the two basic constituents of operational service management. First, formalization of a domain-specific scheduling problem of industry in the necessity to use semantically expressive means for the description of technological limitations and customers' preferences. Second, more and more improving network and multi-core technologies put even in the ordinary personal computers allow to solve scheduling problems at a new technological level. These aspects are reflected in the technologies of ontological knowledge bases and multi-agent systems, related to the area of the distributed artificial intelligence.

According to the provisions of economic theory the mission of the operational management is ensuring a stable process of the primary activity of an organization.

A customer turns to an organization having certain goals and individual preferences. During the dialog with the manager there occurs an assignment of a complex service from a set of standard services rendered by the organization. In other words, a complex service is a reflection of the client's aims on a set of standard services rendered by the organization.

We are set a task of operational service management. So, what are the parameters in this process that we can manage? The controlled variables on the part of the customer in the process of providing services are the ones which describe the desired state of the client. Uncontrollable variables are the beginning and the end of the possible consumption of services, the initial state and physiological features, financial position and status of the client. The effectiveness parameters are the subjective perception of services and preferences.

Controllable variables in operational service management on the part of the service are temporal periods of availability of a service, the information about the mechanism of doing services. Uncontrollable variables are the service location, exploited resources, natural environment, history of doing services, requirements of the technological process and factors of the production environment. The effectiveness parameters are quality criteria and subjective perception of the service by the client.

The proposed conceptual approach to the operational service management is based on a domain-specific complex of models as well as on the chart of an identification type adaptive management. The main acting persons in the service provision process are service user and manager. [2]

The mathematical problem model of assigning services to clients is formulated in terms of the integer programming and is decided by the method of dynamic programming and allows to take into account the economic aspects of providing services and to ground pricing in an organization.

The mathematical model of service scheduling is formulated in terms of game theory with non-contrary interests as a game with the concerted vector of interests with the forbidden situations, and allows to take into account the formalized technological features of the service provision process, individual strategies of customers' behavior, and also behavioral strategies common for all the participants involved in the process.

Presence of non-protuberant area of feasible solutions makes it impossible to solve the given problem by mathematical programming methods.

The method of intellectual decision support is based on:

13. an iterative process of the adaptive planning and making possible administrative decisions;
14. the rules of management decision-making;
15. the agent paradigm of artificial intelligence.

All this provides a theoretical basis for realization of the proposed approach.

Conclusion

In this paper have been described two approaches that take into account uncertain information on clients and their individual preferences. RFMR is based on a scoring approach provide the foundation for the quantification of customer behaviour. The proposed conceptual approach to the operational service management is based on individual preferences in domain-specific complex of models.

Acknowledgements

This investigation is supported by the grants 08-07-00495-a and 09-07-00408-a of Russian Foundation for Basic Research, Academy of Sciences of RB and Russian Humanitarian Scientific Foundation № 09-02-84-201, 2010, President of Russian Federation grant НШ 65497.2010.9 for leading scientific schools

References

1. Client-oriented marketing. The Journal "Laboratory of advertising, marketing and PR", №1 (50), 2007"
2. Hilbert A., Bogdanova D. Information decision support in the operational plant management with the use of technologies of distributed artificial intelligence (on the example of an enterprise of health and beauty industry), CSIT 2008 Proceedings vol.2, pp. 147-154.
3. Miglautsch J.R. "Thoughts on RFM Scoring". The Journal of Database Marketing, 2000; 8: 67-72.
4. Souafi-Bensafi S., Nikravesh M. "Multi-Aggregator Fuzzy Decision Trees: EC-based optimization". In: IEEE International Conference on Industrial Informatics INDIN 2003. Banff, Alberta, Canada, 2003, pp. 379-384.
5. Sakalosky M. "RFM. Part 1". The ClickZ Network. 2002.

Integration of Data Mining Technology and Decision Support System on the Example of Monitoring a Trading Network

K.R.Enikeev
Ufa state aviation technical University
Ufa, Russia
e-mail: kamil_en@mail.ru

G.R.Shakhmametova
Ufa state aviation technical University
Ufa, Russia
e-mail: shakh@ugatu.ac.ru

Abstract[1]

The decision support system for monitoring a trading network, uses semantic network and data mining technologies, is discussed in article.

1. Introduction

Due to the evolution of the data recording and storage methods the expansion of amount of collected and analyzed information is taking place.
In the process of operating a trading network is gathering a huge amount of information. This information can contain the history of the organization, different regularities and often valuable hidden data. Collecting and interpretation of this information help in management decision-making, situation understanding, customers analyses, identification of successful and unsuccessful decisions in the past, forecasting the consequences of the decisions etc. The identified regularities can be of diverse origin.
These problems can be solved via methods of data mining.
Because of large amounts of information, labourous analysis of all possible variants during administrative decision-making the use of Decision Support System is justified. DSS allows work of managers to be facilitated and to increase its efficiency. It considerably accelerates the decision of problems in business. The availability of precisely functioning DSS offers advantages in comparison with competing structures. Because of the offers which are put forward by DSS, new approaches to solving daily and non-standard problems are available.
The article looks at the possibility of integration of data mining and automated DSS, that provides managers with more complete and flexible situation analysis, considering implicit and unknown knowledge, experience and the previous decisions and allows forecasting consequences of decisions.

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

There is a problem of development of such Decision Support System which would use DM and ensure a data storage.

2. Decision support system for monitoring a trading network

The DSS is divided into 3 parts. The first part gathers and stores the data, the second part – DM and the third part – decision-making part which uses the knowledge base.
To achieve these goals the author propose the following general scheme of DSS in monitoring a trading network (Figure 1).
To develop the decision support system for monitoring a trading network it had been used knowledge engineering, expert system technology and data mining technology.
The DM block identifies different regularities and registers them in the knowledge base. Thus KB is constantly updated and specified.

3. Data mining module in DSS

Data mining block can solve problems in customer relationship management [3,4].
Rather than randomly contacting a prospect or customer through a call center or sending mail, a company can concentrate its efforts on prospects that are predicted to have a high likelihood of responding to an offer. More sophisticated methods may be used to optimize resources across campaigns so that one may predict which channel and which offer an individual is most likely to respond to — across all potential offers.
Another example of data mining, often called the market basket analysis, relates to its use in retail sales [5]. If a clothing store records the purchases of customers, a data-mining system could identify those customers who favour silk shirts over cotton ones. Although some explanations of relationships may be difficult, taking advantage of it is easier. The example deals with association rules within transaction-based data. Not all data are transaction based and logical or inexact rules may also be present within a database. In a manufacturing application, an inexact rule may state that 73% of products which have a specific defect or

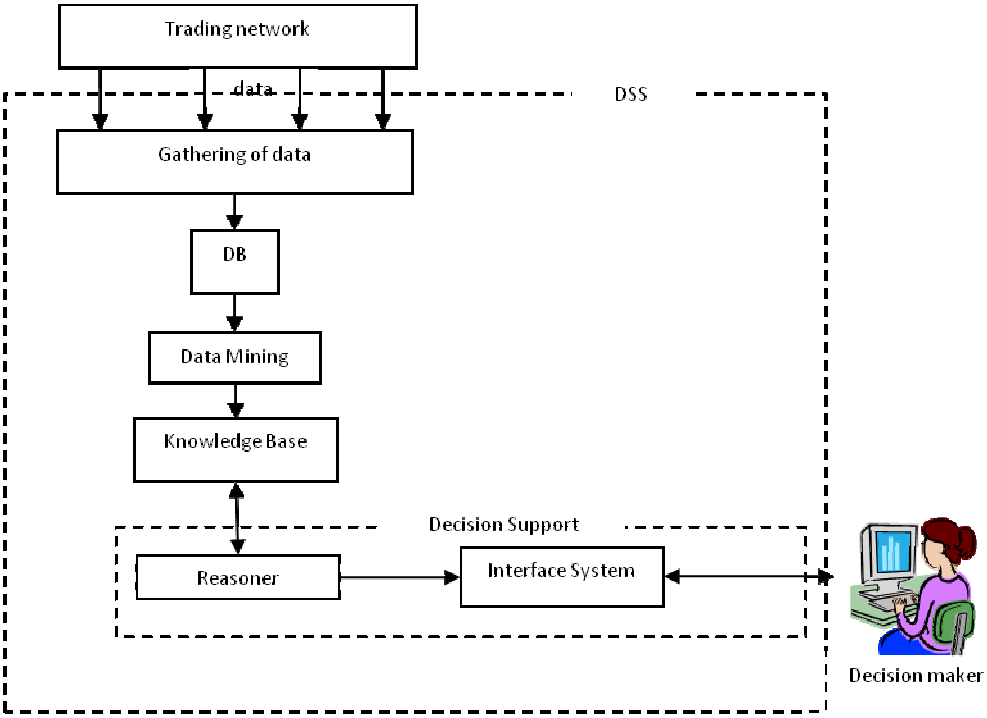


Figure 1. General scheme of DSS

problem will develop a secondary problem within the next six months.

Market basket analysis has also been used to identify the purchase patterns of the Alpha consumer. Alpha Consumers are people that play a key roles in connecting with the concept behind a product, then adopting that product, and finally validating it for the rest of society. Analyzing the data collected on these type of users has allowed companies to predict future buying trends and forecast supply demands.

Apriority is a classic algorithm for learning association rules. Apriority is designed to operate on databases containing transactions, and was used for searching the regularities in purchase of food in one of the suchi restaurant.

As is common in association rule mining, given a set of item sets (for instance, sets of retail transactions, each listing individual items purchased), the algorithm attempts to find subsets which are common to at least a minimum number C of the item sets. Apriority uses a "bottom up" approach, where frequent subsets are extended one item at a time (a step known as candidate generation), and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found.

Apriority uses breadth-first search and a tree structure to count candidate item sets efficiently. It generates candidate item sets of length k from item sets of length $k - 1$. Then it prunes the candidates which have an infrequent sub pattern. According to the downward closure lemma, the candidate set contains all frequent k-length item sets. After that, it scans the transaction database to determine frequent item sets among the candidates.

Apriority, while historically significant, suffers from a number of inefficiencies or trade-offs, which have spawned other algorithms. Candidate generation generates large numbers of subsets (the algorithm attempts to load up the candidate set with as many as possible before each scan). Bottom-up subset exploration (essentially a breadth-first traversal of the subset lattice) finds any maximal subset S only after all $2^{|S|} - 1$ of its proper subsets.

The software implementation of the data mining system to forecast finance indicators of the company is by means of the analytical platform Deductor 5.0.

The results of work are shown on Figure 2.

Results show patterns of purchases of goods.

4. Knowledge base in DSS

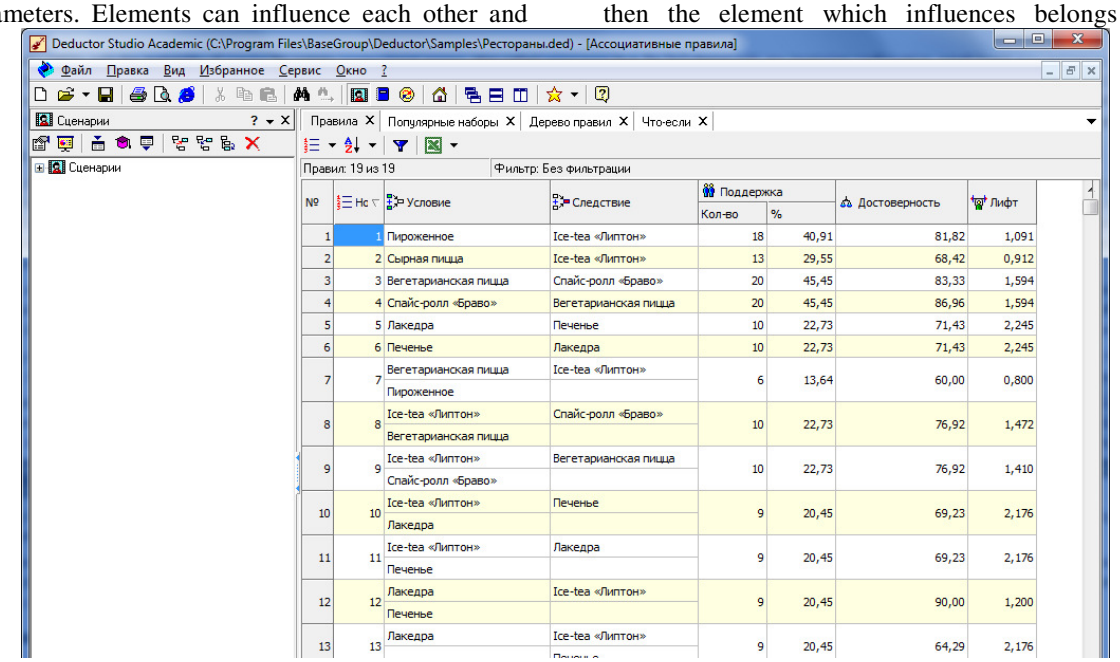
As a form of knowledge presentation the semantic network has been used.

Knowledge base is divided into two parts: the part serving the decision-making logics and a part, describing the subject domain. The other part describing the subject domain is generated according to the ontology structure of decision-making logics.

In Figure 3 there is a graphic representation of the ontology logic of decision-making and the informational model.

The developed knowledge base is a system consisting of elements. Elements have a certain condition – a vector

of parameters. Elements can influence each other and then the element which influences belongs to the



№	Нс	Условие	Следствие	Поддержка Кол-во	Поддержка %	Достоверность	Лифт
1	1	Пироженное	Ice-tea «Липтон»	18	40,91	81,82	1,091
2	2	Сырная пицца	Ice-tea «Липтон»	13	29,55	68,42	0,912
3	3	Вегетарианская пицца	Спайс-ролл «Браво»	20	45,45	83,33	1,594
4	4	Спайс-ролл «Браво»	Вегетарианская пицца	20	45,45	86,96	1,594
5	5	Лакедра	Печенье	10	22,73	71,43	2,245
6	6	Печенье	Лакедра	10	22,73	71,43	2,245
7	7	Вегетарианская пицца Пироженное	Ice-tea «Липтон»	6	13,64	60,00	0,800
8	8	Ice-tea «Липтон» Вегетарианская пицца	Спайс-ролл «Браво»	10	22,73	76,92	1,472
9	9	Ice-tea «Липтон» Спайс-ролл «Браво»	Вегетарианская пицца	10	22,73	76,92	1,410
10	10	Ice-tea «Липтон» Лакедра	Печенье	9	20,45	69,23	2,176
11	11	Ice-tea «Липтон» Печенье	Лакедра	9	20,45	69,23	2,176
12	12	Лакедра Печенье	Ice-tea «Липтон»	9	20,45	90,00	1,200
13	13	Лакедра	Ice-tea «Липтон»	9	20,45	64,29	2,176

Figure 2. Results of Data Mining

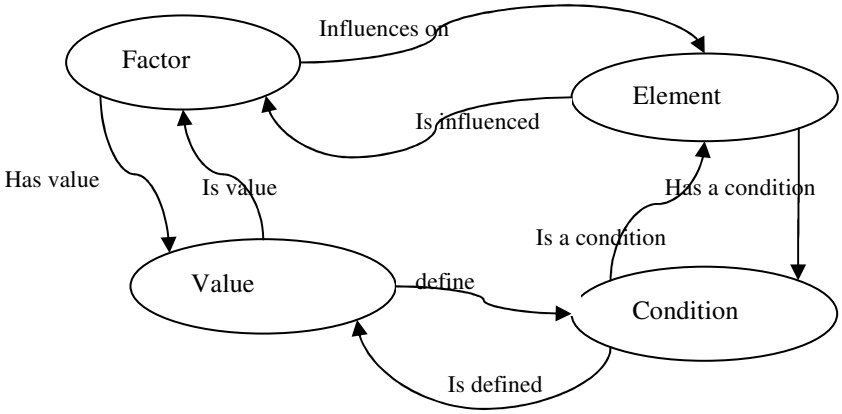


Figure 3. Ontology structure

factor" class , and its condition is called "value". The condition of the dependent element is defined by value of the factors.

The knowledge base is executed in the ontology presentation OWL language, and functional dependences can be transitive, reflective, symmetric and thus by means of the knowledge base it is possible to build the chains of dependences for finding sequences of actions. factor" class , and its condition is called "value". The condition of the dependent element is defined by value of the factors.

The knowledge base is executed in the ontology presentation OWL language, and functional dependences can be transitive, reflective, symmetric and thus by means of the knowledge base it is possible to build the chains of dependences for finding sequences of actions.

Example of part of ontology describing the subject domain is shown in Figure 4.

By means of the data base it is possible to identify the specific type and significances of dependences what allows the program to make concrete calculations.

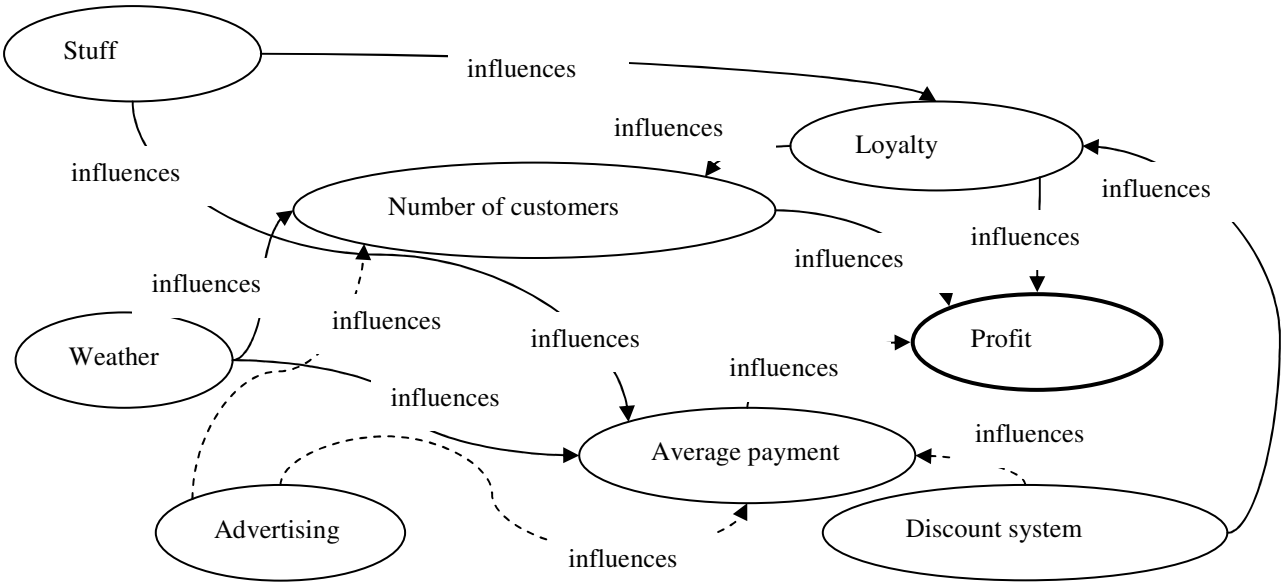


Figure 4. Subject ontology example

5. Conclusion

Block of decision support interacts with user, makes conclusions on the knowledge base. This block gives questions to user in order to find best decision, and makes calculations of indicators of working trading network.

Acknowledgement

This research has been partially supported by grants 09-07-00408-a (Russian Foundation for Basic Research), 09-02-84-201 (Russian Humanitarian Scientific Fund), research work “Research of intellectual technologies for decision-making support and control of complex social and economic objects” (Ministry of Education of the Russian Federation) and by grant of President of Russian Federation № 65497.2010.9 for Leading Scientific Schools.

References

1. Jackson P. "Introduction to expert systems". Williams, Moscow, 2001.
2. Popov E.V. "Statistic and dynamic expert systems". Finance and statistics, Moscow, 1996.
3. Barsegyan A.A., Kupriyanov M.S., Stepanenko V.V., Kholod I.I. "Methods and models of data analysis. OLAP and Data Mining". BHV-Petersburg, Saint-Petersburg, 2004
4. Duk V., Samoilenko A. "Data Mining". Peter, Saint-Petersburg, 2001.
5. John Wang, «Encyclopedia of Data Warehousing and Mining», Montclair State University, USA, 2001

Knowledge Discovery in Databases and its Application in Manufacturing

Michal Kebisek
Faculty of Materials Science and Technology
Slovak University of Technology
Trnava, Slovakia
e-mail: michal.kebisek@stuba.sk

Peter Schreiber
Faculty of Materials Science and Technology
Slovak University of Technology
Trnava, Slovakia
e-mail: peter.schreiber@stuba.sk

Igor Halenar
Faculty of Materials Science and Technology
Slovak University of Technology
Trnava, Slovakia
e-mail: igor.halenar@stuba.sk

Abstract

Today we are witnesses of a continuous development of information technology. It grows exponentially amount of generated and stored data that is available. This often doesn't bring the better knowing but mainly contributes to the disorientation and inability to decide objectively. The excess of data which shall allow to the responsible operator doing of qualified decisions, doesn't lead to understanding of the situation, but often to disorientation and a time stress. Available data still doesn't mean anything to know. Information system filled with data even though has its meaning and form the basis of the "memory" of the organization. The areas having the ability to generate and store large quantities of data have not only areas such as marketing, government, or medicine, but this trend begins to observe in industrial areas increasingly. At the designing of new production lines and equipment is calculated with the possibility of collecting and storing business data from the manufacturing process. Older equipment and lines are often adjusted in order to add this option. It is not necessary to collect such data only, but it is necessary to work with these data properly. Analysts need to obtain information to be able to model objects, anticipate trends and to enable the responsible managers to make appropriate decisions. Classic methods based on SQL, don't enough for it already. For that reason it is the effort to put into this process the methods of knowledge discovery, whose task is to extract a new, valid and potentially useful knowledge from large volumes of available data.
In our article we'll discuss the possibility of the application process of knowledge discovering in databases in the industrial area. We will focus on the initial phases of this process mainly and the problems that can occur during them.

Key words: Knowledge Discovery in Databases, Data Mining.

1. Introduction

For the past few years is typical an explosive increasing possibilities of generating and storing data. In the different spheres, like a science, industry, business, government is data stored and held in the large capacity. Every our telephone call, shopping, internet access or visit a doctor means a large capacity generating data, which are entering in the transaction system of different organizations. That is why organizations obtain potential riches, which creates the bases their so–called "memory". This generates an acute requirement to birth of a new generation procedures and tools for support of discovery usable information and knowledge from all the time faster expand scale electronic data. A lot of organizations cumulate data into their databases, and what is truly needed, are information. The vision about an achievement of information companies, alternatively least utilization of strategic strength deposited in data sources, requests not only new tools, new methods, but mainly new style of thinking. However, problem isn't only in the elaborating of new models. In a matter of fact idea about that, look like neither obtain information about objects, their behavior, requirements, secretion dependences and moreover.

2. Knowledge Discovery in Databases

Information stored in database system is able to divide into two categories. The first category contains the type of information for which was designed database and is about for themselves used. Useful information is

included into specific values of attributes and their combination with regard to relation given database's system. User with queries maniples by the query language, e.g. Structured Query Language so called SQL. The result of query is the set of data incident, where everyone is the mirror of real instance at the outside world. If the database is truthful reflection of reality, results are queries also truthful. With this deductive access method to databases is possible to derive only logical implication information stored into databases. [3]

The second category describes information that is not at first sight visible neither from the attribute values and neither from structure of databases. These information are hidden by the score of database records and it is able to acquire them by the inductive derive process from attribute values. In databases are sought regularities, i.e. combination values specific attributes, common and typical for subset (classes) facts stored into the databases. Results are presented generally as logical formulations expressive conditions decomposition stored data to the individual classes. This information already has not only descriptive character, however indicate specific knowledge about domain, and they can contain fragment of predictions. Results acquired on present–day data are predicting something about on–coming data which still only come into the databases. Derived knowledge needn't to be necessarily truthful in the real world, every consequential formulation it shall have his probability, generally little look like 1. [2]

The requirement finds preferably all interesting information from databases and insufficiency existing procedure, calls into existence establishing a new science discipline Knowledge Discovery in Databases from the ends eightieth years. There are also used designations: information harvesting, data archaeology, knowledge extraction or data mining.

The term Knowledge Discovery in Databases was first–times used on first conference about KDD in 1989 on emphasis the fact, that the objective commodity of data investigations is knowledge. Look like synonym to this term sometimes used term Data Mining. With reference on 1st international Knowledge Discovering and Data Mining in Montreal in 1995 conventus this term use only for one phase of discovering in KDD process.

Knowledge Discovery in Databases definition: KDD non–trivial process identifications valid, recent, till undisclosed, potential usable and well comprehensible knowledge from data. [5]

On the term data we understand the set of ordered facts (e.g. items in database), knowledge's expression in any language describing subset of data alternatively model applied on this subset. That is why we understand on the thereby term discovering knowledge, planning and model specification investigation reality so that describes to data which best, finding description data structure and creation data description on advanced.

Knowledge Discovery in Databases is the process that includes a lot of individual partial steps that take in data preparation, research patterns, verification, testing and refinement receiving knowledge, everything repeated in a many iterations. On the term non–triviality we understand that individual steps include further partial tasks. Validity isn't known absolute, however with regard to any measure reliability. [4] The objective is addition receives information, their ability metamorphoses on any new knowledge. Comprehensibility is often absent.

The whole process of KDD is not possible to automate fully. The effort for full automate of this process leads to achievement of non–optimal results of real valuable knowledge. From this reason is necessary assistance of human operator who determines about choosing suitable operations, algorithms and their parameters within particular steps of process KDD.

Knowledge Discovery in Databases has developed and develops on frontier several research areas look like machine learning, databases, statistics, artificial intelligence, knowledge engineering, optimalization of difficult capacity calculations, image representation of information etc. [1] The basic objective of all these areas is acquisition knowledge from low–level representation in context of a great data volume.

The Knowledge Discovery in Databases process is interactive and iterative, involving numerous steps with many decisions made by the user. Main KDD process is usually divided on these steps:

1. Problem definition – designation of the problem, which is necessary solved by the KDD process.
2. Data selection – data are selected or segmented following some criteria in order to select relevant data set for others applications.
3. Cleaning and preprocessing – stands modification data format and data cleaning, when some data are removal because they are not necessary alternatively they would defend to effective interpretation of query.
4. Transformation – the transformation of reformatted and cleaned data that can be enriched at a further attributes, e.g. from external source and thereby expand applicability not only for used data however for receives results.
5. Data mining – searching of interesting knowledge, those form depends on choosing data mining methods and can have a form of classification rules or trees, functional dependencies, logical rules etc. Results of this step are considerably depends on previous steps.
6. Interpretation and reporting – system identification models are translated look like knowledge that can be used for decision support.
7. Application of discovered knowledge – application of discovered knowledge on the real problem.

The KDD process can involve significant iteration and can contain loops between some two steps. [2]

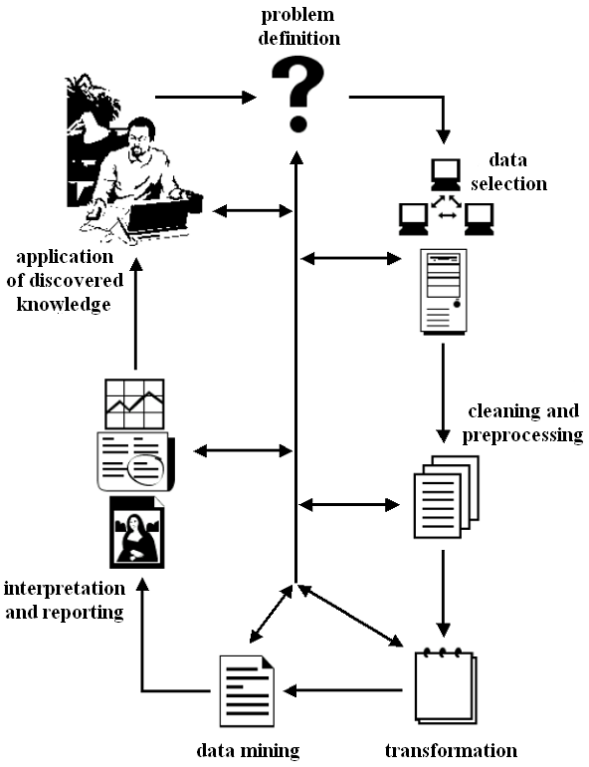


Figure 1. Process of Knowledge Discovery in Databases

3. Data selection, cleaning and preprocessing

In the early stages of the process of Knowledge Discovery in Databases it is necessary to collect and preprocess data with we will work in this process in an appropriate way. During applying of this process in an industrial area should be adjusted not only data stored in databases but it is necessary to ensure the proper data collection from the production system.

3.1. Modification of the production system in order to collect the required data

To work with data from the production system requires that the production system can produce, store and make available appropriate data about their activities. These properties do not provide all the commonly used production equipments. If the production device does not allow collecting working data than can be replenished with new module that will collect production data and allows them to store. Simpler devices that can collect and preprocess at least basic information e.g. bar code readers, industrial cameras, or optical gates. For ensuring of the possibility of reading and storing large amounts of appropriate data can be used for example intelligent conveyors or pallet, RFID (Radio Frequency Identification) code sensors, PLC (Programmable Logic Controller) etc.

For some production devices cannot be applied the auxiliary modules for the collection and storage of working data an here is necessary to ensure the entry of production data otherwise e.g. for the employees who serve the production device.

Newer production devices such as NC (Numerical Control) and CNC (Computer Numerical Controlled) devices, service robots and manipulators, devices with OPC (OLE for Process Control) architecture etc., have already in most cases, custom modules for the collection and storage production data.

In this case, it is only necessary to modify which data and in what format they will be collected and how they are made available for storing to the production database.

3.2. Cleaning and preprocessing of data obtained from the business database

The data discovered from the business database is necessary to clean and preprocess in an appropriate way. If all the necessary data are stored in one database, their data modifying is in this case relatively simple. They generally consist of data format modification, changing the data specifications, value unification of some parameters, value aggregation of specific parameters etc.

If data are stored in different data store which maybe relational or file database, text files etc., it is necessary except of the basic modification, such as data format modification, changing the data specifications etc., to perform further modification. It is required to identify explicitly the common items in different data source, to unify their name convention and to define the correct rules for their exact identification.

Also it is necessary to remove sparse and non-consist data, to identify outlier data that are in their value significantly differ from another data. At the identification of outlier data it is necessary to deal with their thorough investigation. There may be the random data, those extreme values are caused by error during data typing, but it may be important value, caused by non-standard order. It is therefore necessary to identify the origin of those values and to decide whether to include these values in the current data set or to remove them from current data set.

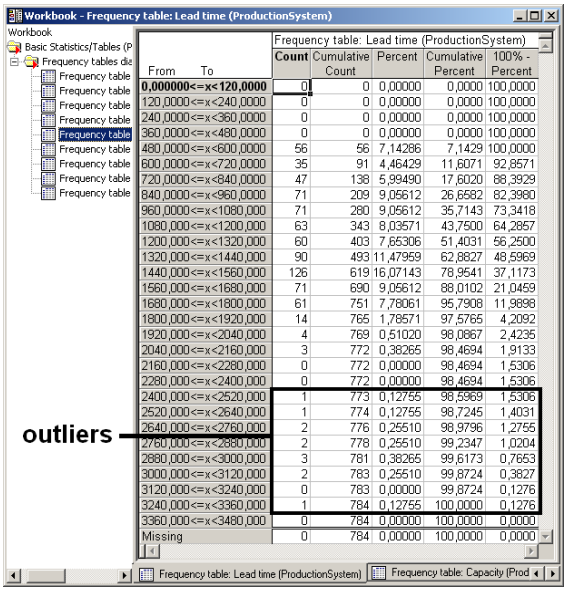


Figure 2. Outliers

The result of this step impacts the quality of next steps directly that will follow in the process.

3.3. Cleaning and preprocessing of data obtained from the production system

During the cleaning and preprocessing of the data obtained from the production system it is also necessary to realize similar operations like in the previous step.

The problems may arise with the lack of such data. Data can miss due the failure of the connection between production devices and the database or can miss from the production devices that are not possible to collect data etc. It should be an appropriate way to solve this problem. The generating of similar data can resolve the problem of missing data from the temporary interruption of the connection between the production device and database. Some data mining tools contain the module that generates missing data on the base of production data which production device created during the similar operations that should preferably equal production device and process setting etc. Another possibility is to refill missing data by the mean values, most probably values, constants or if miss only small amount of data, it is possible to refill it manually.

If it is not possible to refill the missing data with equally appropriate data, it is necessary to make provision for this fact during the data mining process and especially during the evaluation of the discovered results.

3.4. Data transformation

During this step of process are merged the data obtained from production system and data obtained from business databases. In previous step of process must be unified the names, formats etc. If this is not performed it is necessary to perform them now. It is important to identify data correctly not only during the data mining process, but also during the evaluation of the discovered results. It is necessary to remove redundant data that may occur when the data are merged from the production system and the data from the business databases.

If the suggested data mining method or technique does not have to work with the whole data set, it is in this step possible choice the appropriate data subset that the data mining method or technique will operate with. Also it is possible to reduce used data set regarding to time interval such as data from the last week, month, year etc.

It is necessary to provide the transformation depending on defined purposes and predicted data mining methods and techniques that will be used during this process and also depending on used data mining tool.

4. Conclusion

The process of Knowledge Discovery in Databases needs appropriately adjusted data from which obtains valid, recent, till undisclosed, potential usable and well comprehensible knowledge. For this purpose it is necessary to collect appropriate data from the production system. Thus collected data and data stored in business databases can be used as input data set in the process of Knowledge Discovery in Databases.

References

1. Elias M., Palaj J., Durci M. “E-archiving of paper engineering documentation”. In: *International Doctoral Seminar 2008 Smolenice*, AlumniPress, Trnava, Slovakia, 2008, pp. 60–64.
2. Fayyad U. M., Piatetski–Shapiro G., Smyth P. “From Data Mining to Knowledge Discovery: An Overview”. In: *Advances in Knowledge Discovery and Data Mining*, AAAI Press/The MIT Press, Menlo Park USA, 1996, pp. 1–30.
3. Halenar R. “Phase of data extraction and processing from OLTP systems for use in DWH”. In: *Journal of Information Technologies*, Trnava, Slovakia, 2010, pp. 18–24.
4. Tanuska P., Moravcik O., Vazan P. “The base testing activities proposal”. In: *Annals of DAAAM and Proceedings of DAAAM Symposium*, Vol. 20, No. 1 Annals of DAAAM for 2009 & Proceedings of the 20th international DAAAM symposium “Intelligent manufacturing & automation: Focus on theory, practice and education”, DAAAM International Vienna, Austria, 2009, pp. 371–372.
5. Trnka A. “Proposal of using selected data mining methods in Six Sigma methodology”. In: *Journal of International Scientific Publication*, Materials, Methods & Technologies, Vol. 4, No. 2, Trnava, Slovakia, 2010, pp. 409–417.

Intellectual Drilling Colums and Information Technologies in Oil Extracting

A.V. Shchenyatsky
Izhevsk State Technical University
Izhevsk, Russia
e-mail:fmim@ustu.ru

I.A. Galikeev
OOO "NPP Horizon"
Izhevsk, Russia
e-mail: mail@horizon-drill.ru

E.S. Chukhlantsev
Izhevsk State Technical University
Izhevsk, Russia
e-mail:OpenGL111@ya.ru

Abstract[1]

The chisel column with two-planimetric hydraulic deviator, further on called an intellectual chisel column, which can deviate in the course of drilling on the set corner under corresponding program, is developed. By means of an information technology and the analysis of soil the COMPUTER develops the trajectory of the column movement.

1. Introduction

In the whole history of oil and gas industry there has been an evolution aimed at increasing the efficiency and growth of the oil recovery factor of the deposit. At the initial stage of the field development, oil was extracted from vertical wells (one-dimension) (Fig. 1, a) [1], their number and location were not taken into consideration, only the product availability was important. With the industry monopolization, the efficiency of investments was getting important. As a result, there was the transfer to system development of fields with vertical wells. The stage of field development with vertical drilling can be characterized as the stage of spot completion of productive deposits with one-dimension wells (Fig.1, a).

2. History of oil production

Together with the simplicity and a number of technological advantages, vertical wells did not allow developing the fields under artificial and natural obstacles. To solve this problem, the technology of drilling inclined wells was developed. The drilling technology and well exploitation reached a new level. However, the deposit completion was still a spot one, though due to the initial vertical drilling, the wells became two-dimension [2] (Fig. 1, b).

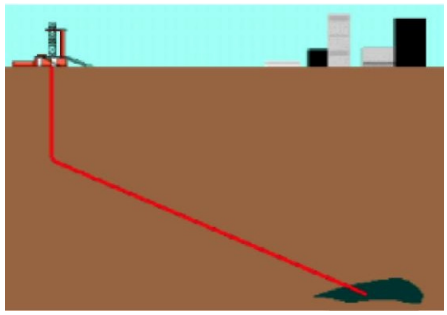


Figure 1– a – one-dimension well

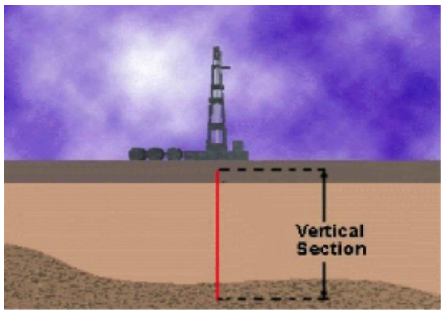


Figure 1– b – two-dimension well

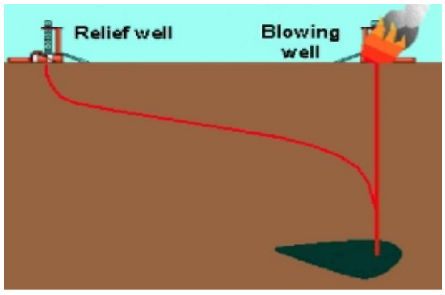


Figure 1– c – three-dimension well

Further development of building technology of inclined wells, allowed developing the drilling of horizontal wells, thus becoming a new stage in developing the efficiency of oil and gas field exploitation (Fig. 1, c). The main difference between the horizontal and vertical or inclined wells is that not spot but linear field

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

completion is applied, thus significantly increasing the field coverage, surface filtration and oil recovery factor (ORF), decreasing the environmental impact, providing the significant growth of the efficiency of capital investments. The necessity of coupling with the initial vertical region results in the need to design 3D well.

Together with the changes in field development techniques there was an evolution of technical means for their implementation. Vertical wells were drilled either rotary with the drive from the rotor or cable-tool. However in the middle of XX century, mainly in the USSR, the drilling of wells with turbo drills – multi-stage hydraulic turbines providing the rotary motion of rock-destruction tool only with hydraulic energy of drilling liquid without rotating the whole column was developed [3]. Together with increasing the drilling speed, the technology also allowed solving the issue of drilling inclined wells. In the middle of XX century the largest oil fields in Western Siberia were discovered, where the major problem was to provide the proper coverage in swampy areas.

Initially the inclined wells were drilled with turbo drills with adapter subs with the misalignment of threaded connections. As a rule, relatively large turbo drills (over 10 meters) and arrangement rigidity did not allow using adapter subs with large misalignment. As a result, only the inclined wells with great curvature radius (over 350 meters) had to be drilled. However, when shifting to spindle turbo drives, the technical means for inclined drilling also change. Despite the fact that there was no significant change in the column rigidity, there appeared the possibility to bring the oblique adapter sub to the drilling bit and install it between the spindle and working part of turbo drill [4], thus 1.5 – 1.8 times decreasing the curvature radius with consequent increase in maximum inclination angles and displacements. Such change in the technology allowed sufficiently increasing the field coverage from one multiple-well platform. Turbo drills have higher rigidity due to their sizes, as well as sufficient weight, when reaching the inclination angles over 45-50° there arise problems with bringing the load to bottom hole. High frequency of motor shaft rotation does not allow applying adapter subs with great misalignment. The constant misalignment does not give the possibility to promptly control the trajectory, etc. [5].

In 1970-s screw bottom-hole motors with significantly smaller sizes were developed, which allowed providing the torques several times exceeding the torques of turbo drills. Together with bottom-hole systems of parameter control, screw bottom-hole motors made it possible to drill wells with curvature radii up to 20 meters with conditionally unrestricted inclination angles. The less motor revolutions allowed installing a special device between the operating steam and spindle that changes the inclination and, consequently, the well curvature radius. At the same time, there were some shortcomings left when applying these steering tools: change in the curvature radius is possible only when lifting the drilling tool and changing the misalignment, the curvature parameters are got with fixed drilling tools, thus decreasing the quality of well walls, the speed of drilled formation lifting, the load bringing to the bottom-hole.

In the late 1980-s a new tendency in the development of inclined and horizontal drilling appeared – controlled rotor drilling [6]. When applying this technology a number of record-breaking wells with well head displacement over 10,000 meters have been drilled. The advantages of this method: improved quality of well bore, improved returns due to the flow artificial turbulence, better load bringing to the bottom-hole.

At the same time, together with controlled rotor drilling there is a tendency of producing various devices used to change the direction of the column motion (Fig. 2) [7].

The devices providing the column deviation can be divided into 2 groups. The first comprises all assemblies containing special fracture elements providing the curvature of the column longitudinal axis. The second – different types of column deflectors on the predetermined inclination angle (Fig. 2). In this case, the relative position of the column axis and well axis change.

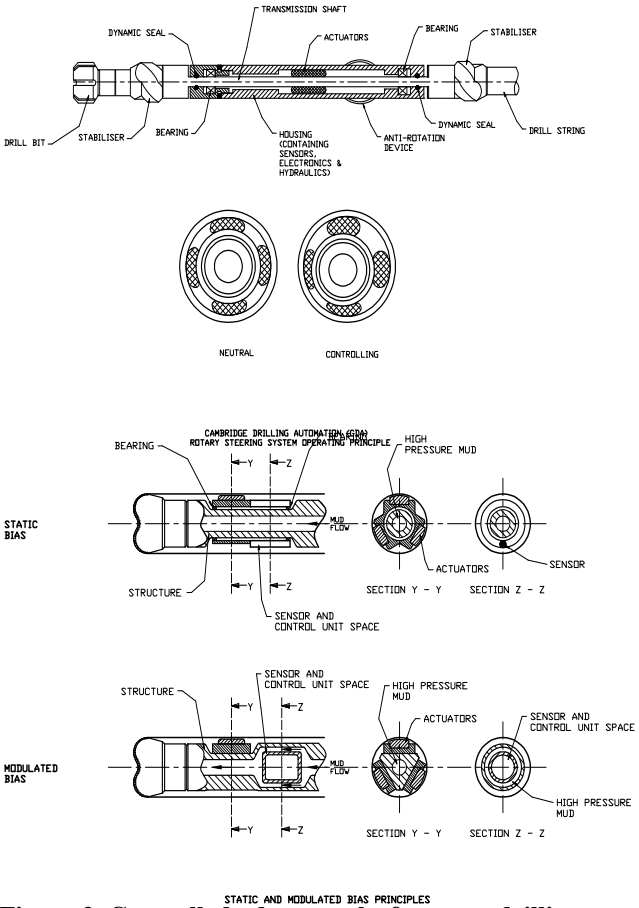


Figure 2. Controlled adapter subs for rotor drilling

3. Analysis of contemporary designs of drill stem

The analysis of column modern structures and technologies tested in practice revealed that there is considerably better access to main deposits of hydrocarbon raw materials concentrated in hard-to-reach areas – shelves, swampy regions, tundra, permafrost areas. Besides, it is possible to produce difficult to recover stores – residual stores of old fields and deposits with low collector properties, and high-viscous oil and tar. In all the foregoing situations it is necessary to accurately drill the well with complex trajectory, as well as to many times change the inclination angle to drill through the complicated areas, which currently requires multiple hoisting of the column. Obviously we need to drill such wells with robotized systems excluding extra technological operations.

Rough drilling conditions, impact loads, alternate conducting and dielectric layers, certain properties of drill fluids give the possibility to apply complex robotized bottom-hole systems in a very limited way. The impossibility to carefully forecast the muck properties in the drilling process, and, moreover, to produce an autonomous analytical bottom-hole complex, predetermine the necessity to use a robotized complex controlled from the surface.

Technological parameters and drilling conditions allow using hydraulic, ultra low frequency electromagnetic or cable communication channel [7]. The shortcomings of the hydraulic communication channel are: low speed of data transfer, dependence on the uniformity of drilling liquid stream, but when drilling productive deposits in a sparing mode, the possibility of sending signals is unrealistic. The electromagnetic channel depends on electrical properties of the muck and drilling liquid, which require the use of ultra-low frequencies significantly decreasing the speed of data transfer. Cable channel requires a lot of connections resulting in bridging and loss of data.

4. Conclusion

The possible solution can be the use of traditional drilling rigs produced as robotized systems with double hydraulic and electromagnetic communication channel, duplex in both channels. The use of rigs with "continuous pipes" assembled, gives the possibility to effectively apply a cable communication channel. The cable communication channel will also be effective and in demand when drilling with bottom-hole electrical motors (electrical drills) that have a number of technological advantages when compared with other drilling techniques.

The task to control the trajectory during drilling means to control the change in rotation axis of the drilling tool against the column by value, as well as by the direction against the apsidal plane. The robotized column with bottom-hole motor is controlled from the surface without any difficulty, but, as mentioned before, there are several limitations. The possibility to combine the rotation of the column and control of the drilling direction is available only with the fixed body element of the column that provides the torque transfer in the required direction. As the developed and trialed models demonstrated, the rotation axis can be effectively changed in two ways: producing the misalignment in the foregoing fixed body element similarly to the misalignment in bottom-hole motors; or decentralizing the fixed column body.

The analysis and approbation of the drilling columns of various models and design demonstrate that the use of a number of decentralizers as deflectors from the given movement direction results in the new technique of controlled drilling, without column hoists. When the trial model, and, further, the manufactured column possess the required technical and durability characteristics, their application in practice will allow significantly increasing the drilling speed of directed and multihole wells, decreasing the process costs.

The Department of Mechatronic Systems of Izhevsk State Technical University and NPO "Gorizont" are working on entirely new scheme of the drilling column with two-contour hydraulic deflector of controlled drilling providing the change in the drilling angle from 0º up to 3º with the error 0.08º – 0.15º depending on the performance level.

References

1. Tagirov K.M., Nifantov V.I. Well drilling and completion of oil and gas layers on drawdown. – M.: LLC "Nedra-Biznestsentr", 2003.
2. Kalinin A.G., Levitsky A.Z., Messer A.G., Solovyov N.V. Manual on well drilling technology of liquid and gaseous minerals: Reference-book / Edited by Kalinin A.G. – M.: LLC "Nedra-Biznestsentr", 2001.
3. Bulatov A.I., Makarenko P.P., Proselkov Yu.M. Drilling liquids and grouting mortars: Students' aid for HEIs. – M.: OJSC "Publishing house "Nedra", 1999. – 424 p.
4. Basarygin Yu.M., Bulatov A.I., Proselkov Yu.M. Drilling technology of oil and gas wells: Students' aid for HEIs. – M.: LLC "Nedra-Biznestsentr", 2001. – 679 p.
5. Boldenko D.F., Boldenko F.D., Gnoevykh A.N. Screw bottom-hole motors. – M.: Nedra, 1999.
6. Bulatov A.I., Kachmar Yu.D., Makarenko P.P., Yaremiychuk R.S. Well drilling: Reference-book. – M.: Nedra, 1999.
7. Drilling equipment: Reference-book in 2 volumes. – M.: Nedra, 2000. – V. 1. – 269 p.

Increase of Automatic Temperature Control Effectiveness of Thermal Objects Based on Application of Pulse Energy Method

Ivan Davydov
Votkinsky Branch of Izhevsk State Technical University
Votkinsk, Russia
e-mail: davyd85@mail.ru

Alexander Shelpyakov
Votkinsky Branch of Izhevsk State Technical University
Votkinsk, Russia
e-mail: alshelp@mail.ru

Abstract

Based on the energy representation the solution of the regulation proposed pulsed power method of regulating the output parameter object. Based on research results it is developed an algorithm of adaptive control system using a pulsed energy method with elements of fuzzy logic to control the inertial objects with delay. The results of a new automatic control system showed the effectiveness of management in comparison with existing systems.

Introduction

The main tasks of the automatic control are to improve the accuracy of the regulation and reduction of the duration of the transition. The process of managing thermal inertia of objects is characterized by the growth of inertial controlled quantity (temperature), after disabling the control action. That's why the management of thermal processing equipment has long transients and large amplitude of the overshoot due to inertia as a result, including the use of positional and PID laws of control.
The objectives of improving the efficiency of automatic control are: improving accuracy, reducing the duration of transients and the sustainability of the regulatory process.

Publishing

The regulatory process of the parametric state of the object, in essence, is an energy process in which the flow of energy from the system of regulation comes to an object, changing its energy and, consequently, the parametric condition [1].
To control the state of the object the energy flow can be arranged either in the form of individual pulses of constant amplitude, but with a variable duty cycle, which appears as the ratio of the pulse duration to the period of repetition of these impulses. In accordance with the energy representation of the regulatory process can be assumed that for a given parametric object is necessary for it to file a certain flow of energy. Mathematical and physical transient study confirms the linear character of dependence of the steady state of an object on the value supplied to its energy. Fig. 1 shows the dependence of the steady-state value of the output signal Y on the value of the energy flux (power) control action. As the object investigated the thermal object, the output signal Y which is the temperature T (°C) as a control action applied electrical energy in the form of a continuous signal a certain voltage U (V) and current J (A) supplied to the heating element. Flow control is defined as the energy output of the control signal: P=U*J (W).

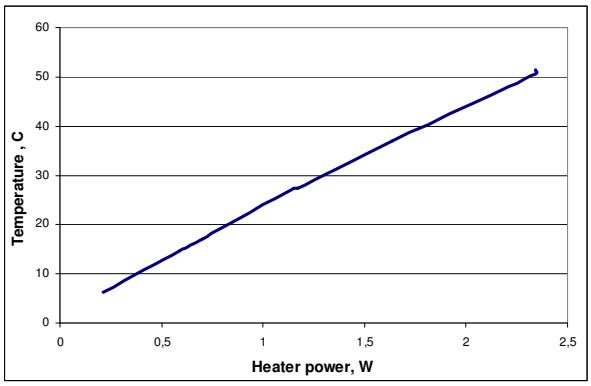


Fig. 1. Dependence of temperature on the heater power

Fig. 2 shows the dependence of the steady-state value of the output signal Y (°C) on the value of intermittency factor, $S=t_{И}/t_{СЛ}$ control signal certain of constant amplitude (U*J).

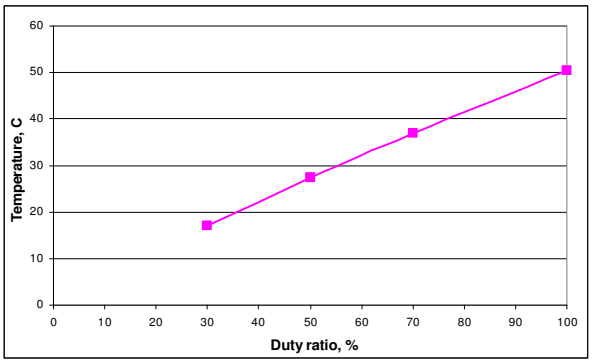


Figure 2. Dependence of temperature on the duty ratio control action

Based on the energy representation the solution of the regulation proposed pulsed power method of regulating the output parameter object [2], which essentially consists of the following. Originally formed by the control action in the form of a pulse peak power and certain duration, ensuring the withdrawal of the object for a given value of the output parameter Y_3 in minimum time. Then, the forming of the control pulses of definite duration $t_И$ with a certain period $t_{СЛ}$ (duty ratio $S=t_И/t_{СЛ}$) to ensure accurate output of the object at a given value of output parameter Y, and maintain it in this state. To ensure the proposed method of regulation should largely determine the two control parameters: the duration of the first pulse, which provides inversion of an object to a new desired state and the value of duty ratio S of the control pulses in order to maintain the object in the given state.

In [3] describes the formation of the initial control pulse.

To determine the duty ratio S, considering the results of presented researches and energetic character of the object in estimated value of S duty ratio control signal, providing the maintenance of an object at a particular value of Y_3 can be defined by the following relationship:

$$S_P = \frac{Y_3 - Y_0}{Y_{MAX} - Y_0} \qquad (1)$$

where Y_3 - given value of the output parameter object (set point), Y_{max} - approximate maximum value of the output signal Y at a constant continuously manage the impact of a given power;

Y_0 - the initial value of the output parameter (ambient temperature).

For ensure accuracy in the regulatory process requires constant adjustment of duty ratio ΔS signal, which is the definition of the current duty ratio by the formula:

$$S = S_P + \Delta S \qquad (2)$$

Value ΔS can be determined based on the following reasoning. When controlling parameter of the object near the coordinates of the Y_3 may be following the situation presented in fig. 3.

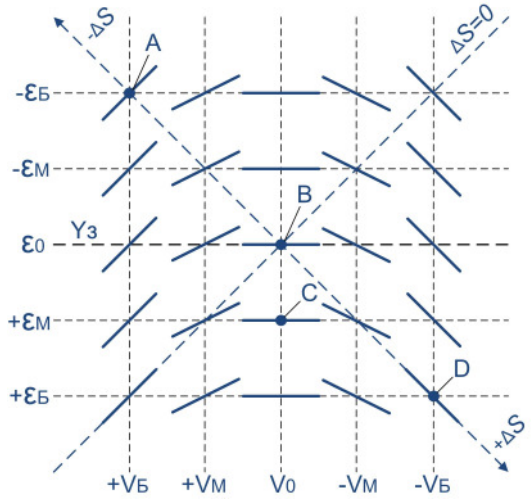


Figure 3. Phase diagram of the regulatory process

The considering process can be divided into two basic values: the error control $\varepsilon=Y_3-Y$ and rate of change of output signal. At fig. 3 the average horizontal dotted line shows the signal value Y_3, solid lines - fragments changes the output signal Y, reflecting the different situations and combinations of error rates ε and V.

The entire range of values can be divided into the following meanings: $+\varepsilon_Б$ - a large positive deviation, $+\varepsilon_M$ - small positive deviation, ε_0 - a zero deviation, $-\varepsilon_M$ - a small negative deviation, $-\varepsilon_Б$ - a large negative deviation, as well as $+V_Б$ - a large positive velocity, $+V_M$ - low positive rate, $+V_0$ - zero rate, $-V_M$ - a small negative rate, $-V_Б$ - a large negative velocity.

Using elements of fuzzy logic ε and V for the specified space variables can take the conventional values ΔS. For example, for the situation (point A) in Fig. 3, there is a large positive velocity ($+V_Б$) and a large negative error ($-\varepsilon_Б$), it is obvious that for this situation, adjust the quantity ΔS should be as negative ($-\Delta S_{MAX}$); for the situation (point B) value ε_0 and V_0, therefore, the value of ΔS=0; for the situation (point C), where $+\varepsilon_M$ and V_0 requires the average positive value adjustments ΔS; for the situation (point D), where $+\varepsilon_Б$ and $-V_Б$ requires the maximum positive adjustment $+\Delta S_{MAX}$. The functional relations between the values ε, V and ΔS shown in the three-dimensional diagram in fig. 4.

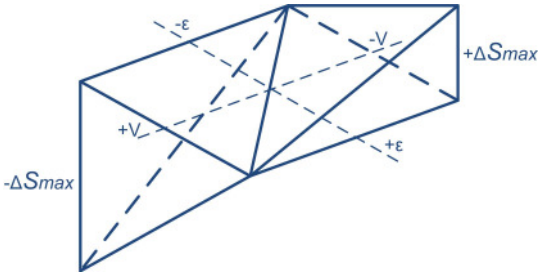


Figure 4. Three-dimensional diagram of the dependence of the duty ratio of the adjustment variables ε and V

The values $-\Delta S_{MAX}$ and $+\Delta S_{MAX}$ can be defined as follows. After determining the calculated duty ratio S_P impulse control is done. Pulse shaping is presented in fig. 5. Then the value $+\Delta S_{MAX}$ is equal to the period of pause to the period of impulse consequence or $+\Delta S_{MAX}=1-S_P$. Then $-\Delta S_{MAX}=S_P$.

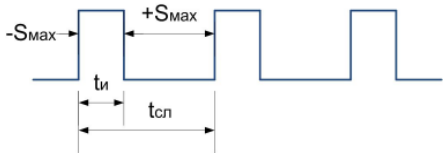


Figure 5. The Determination of values -ΔSmax and +ΔSmax

The values of ΔS can be determined with the following method. The set of values given in the diagram (fig. 6) allows determining the magnitude of the correction ΔS calculated duty ratio S_P in the current state of the regulatory process, i.e. the values of the current error control ε and speed V.

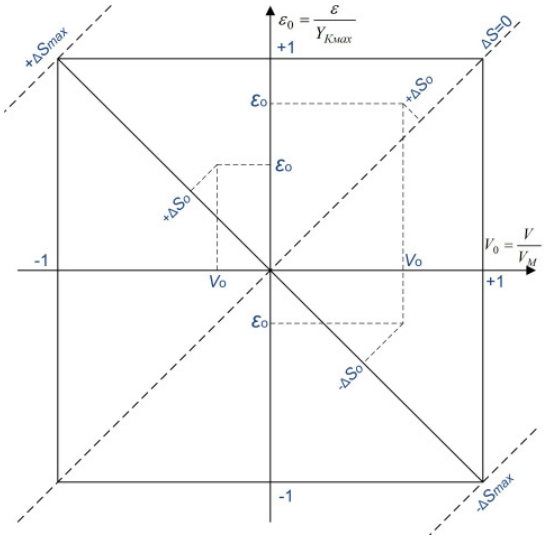


Figure 6. Diagram determines the correct duty cycle control action

To summarize the magnitude of the error of this method of regulation and the rate of temperature change adopted by as relative and are defined by formulas (3) and (4), respectively, which allows the system to take into account the energy state of the object corresponding to Y_3.

$$\varepsilon_0 = \frac{\varepsilon}{Y_{Kmax}} \qquad (3)$$

where Y_{Kmax} is the maximum value of inertia increase of the output parameter of the object after the first single pulse.

$$V_0 = \frac{V}{V_M} \qquad (4)$$

Where V_M - maximum rate of change of the output parameter object in the transition process.

After determining the values of relative errors and the speed is determined by the value of correction of duty ratio ΔS by the formula (5):

$$\Delta S = \frac{\varepsilon_0 - V_0}{2} \qquad (5)$$

Then the duty cycle control signal S is determined by the formula (2).

It should be noticed that the parameters of the control object can be changed in the regulatory process. For example, increasing voltage will increase the maximum furnace temperature, i.e. an increase in the parameter object Y_{MAX}, and the formation of scale in the boiler or lowering the ambient temperature will reduce the value of Y_{MAX}. Therefore, to compensate the changes of the parameters of the control object in certain moments, the correction values of Y_{MAX} is done.

Redefinition of parameter Y_{MAX} is done using a linear dependence on the controlled variable duty cycle. Therefore, the value of Y_{MAX} is determined by the formula (6):

$$Y_{MAX} = \frac{Y - Y_0}{S} \qquad (6)$$

After reordering parameter Y_{MAX} again duty ratio of control is calculated of S_P signal. Then according the current state of control object determines the value of duty cycle correction ΔS and to the appropriate value is determined by the desired duty cycle signal S_P and ΔS, necessary to maintain the controlled variable at the desired set point regulation of Y_3.

Based on research results it is developed an algorithm of adaptive control system (fig. 7) using a pulsed energy method with elements of fuzzy logic to control the inertial objects with delay.

Hardware implementation of the algorithm is allowed to regulate the temperature of thermal objects with sufficient accuracy. Comparative analysis of qualitative characteristics of the transients presented in fig. 8.

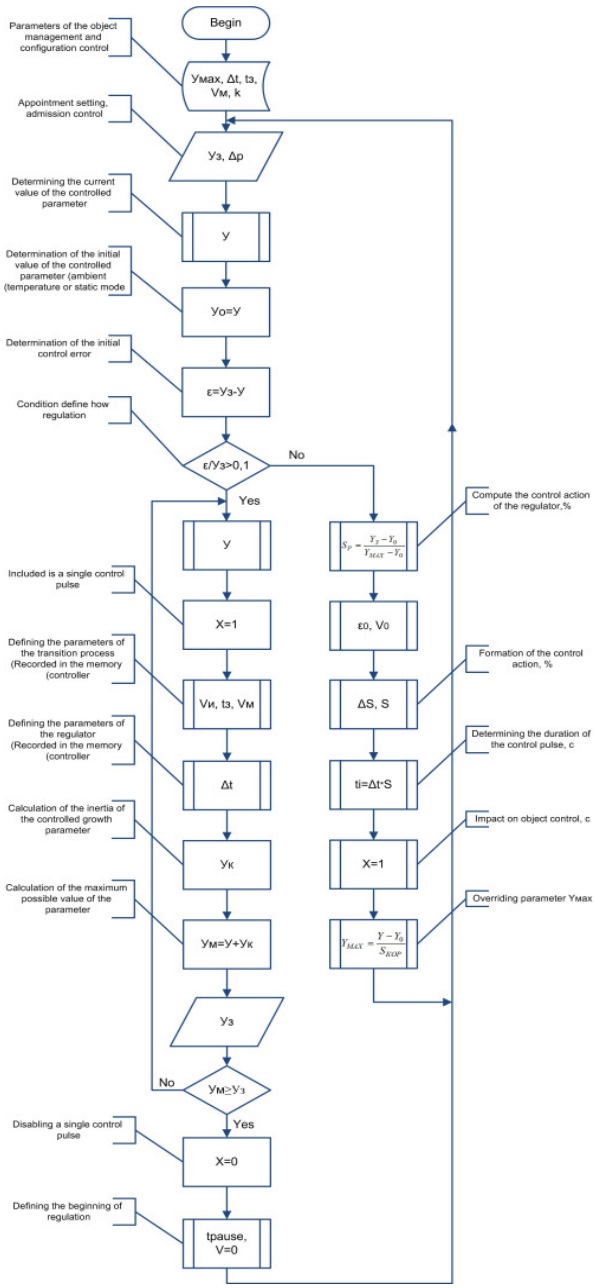


Figure 7. An algorithm for adaptive control using a pulsed energy method with elements of fuzzy logic

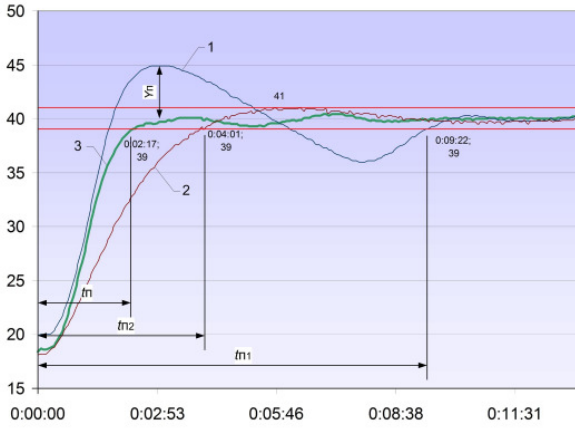


Figure 8. Comparative analysis of qualitative characteristics of transients

When comparing the quality of the regulatory process on real objects had been taken following controls:

-PID-controller, tuned for maximum speed out on the installation - 1;

-PID-controller, tuned to a minimum temperature overshoots - 2;

-Regulator, using an adaptive algorithm with a pulsed energy method - 3.

Analysis of characteristics shows that the regulation of the first regulator is characterized by large amplitude of the overshoot $Y_П$ and prolonged duration of the transition process tп1. The second regulator provided the minimum overshoot, but the duration of the transition process tп2 is still big enough. Application of adaptive algorithm with using of pulse energy method made it possible to ensure the minimum possible duration of the transition process tп without overshoot and with the subsequent satisfactory accuracy regulation.

Conclusion

The results of a new automatic control system showed the effectiveness of management in comparison with existing systems.

It should be noted the merits of the algorithm:

- Adaptation of the controller to the control object. Thus there is a determination of the parameters of the transition process, according to which adjustments are calculated.
- Achievement of adjustable range setting for the minimum possible time,
- Satisfactory quality of control in the process of maintaining the value near the installation by determining the flow of energy in accordance with the adaptation of the controller to the control object (definition Y_{max}, S_p, ΔS and $S_{кор}$).

References

1. Shelpyakov A. “To the question about the energy pulsed method control of the object parameters”. In: Science. Technique. Education: Proceedings of the scientific-methodical conference Votkinsk branch ISTU, Ekaterinburg-Izhevsk, Publishing Economics Institute UB RAS, 2008, pp. 86-91.
2. Shelpyakov A., Davydov I. “Pulse Energy Method of Objects Control”. In: Intelligent systems in manufacturing, Vol. 1. ISTU, Izhevsk, Russia, 2010, pp. 320-323.
3. Davydov I., Shelpyakov A. “Research of parametres of transients at influence on object of a single pulse”. In: Measuring, monitoring and diagnostics, Project, Izhevsk, Russia, 2010, pp. 152-157.

Cross-Platform System of Capturing and Analysis of Corporate Internet Traffic with Artificial Intelligence to Detect Anomalies

Andrey Konin
Izhevsk State Technical University
Izhevsk, Russian Federation
e-mail: andrkonin@mail.ru

Alexander Korobeinikov
Izhevsk State Technical University
Izhevsk, Russian Federation
e-mail: kav33@inbox.ru

Abstract

A system of capturing and analyzing of Internet traffic, designed for Linux and Windows operation systems is presented in this article. The mechanism of analysis of captured Internet traffic through the use of different filters, clustering and subsequent classification of packets to a potentially dangerous and safe with using COBWEB algorithm in WEKA.

1. Introduction

With growing using of the Internet payments, purchases, and money transfers in Global Information Society, increases the number of threats and vulnerabilities posed by unethical users and violators of information security. The number of data packets and various protocols that pass through the network card increases with increasing of speed connections provided by internet companies. For the analysis of each, standard anti-virus requires an unlimited amount of time. If the connection speed is quite high, hence the speed of loading web pages and conduct critical operations has multiplied.

2. Pcap library

To work with the network traffic the library pcap[1] is using. It is implemented for all Windows-platforms (95, 98, ME, NT, 2000, XP, Vista, 7) and for most UNIX-based systems. Incorporating interfaces, libpcap and WinPCap, are different, but, with the help of pcap it is possible to create a portable network analyzer on any system for which pcap library implemented. For example, a network analyzer with support for BPF-Filter (universal decision for almost any task), compile the same on Windows and Linux. Except possible warning, because of mismatch.
Thus, the pcap library allows create programs to capture and analyze network data arriving at the NIC to the PC. An example of software that uses the library Pcap, called Wireshark. Various programs for monitoring and testing network, use this library. It is designed for C-like programming languages, so that programming languages such as Java,. NET and others, use this shell. For Unix-like systems used libpcap library, and for Microsoft Windows NT used WinPcap library. Software network monitoring can use libpcap or WinPcap to capture packets in the network, and (in newer versions) to send packets in the network. Libpcap and WinPcap also support the preservation of captured packets in a file and reading files containing saved packets. Programs based on libpcap or WinPcap, can capture network traffic, analyze it. File, captured by pcap are stored in format, that is understandable for applications using the Pcap.
To work with pcap library requires wpcap.dll (implementation of pcap for Win32), packet.dll (service provider for wpcap.dll) and NDIS-driver kernel level, include-files and export files compiled for wpcap.dll and packet.dll. In version 3.0a there is a small bug that prevents the normal assembly of wpcap.dll and packet.dll for NT/2000/XP. To fix it, you must use the version above 4.0 [2].
To gather information about the available interfaces, packages are available for analysis, using function int pcap_findalldevs (pcap_if_t ** alldevs, char * errbuf). In addition to the network adapters it allows to choose infrared devices, parallel ports and other devices (Fig. 1).

Parameter pcap_if_t ** alldevs - singly linked list.
List of available network devices can be made as follows:

```
pcap_if_t *devlist;
pcap_if_t *d;
if(pcap_findalldevs(&devlist,errbuf)
== -1)
{ handling error here; }
for(d=devlist;d;d=d->next)
{ some code here; }
```

** Alldevs - interface element necessary for communication with the user. If there is no need to store the list in memory, destructor - pcap_freealldevs (pcap_if_t * alldevp).

As for interface names, their names correspond to the Windows code devices, their types and other options, in UNIX displays the names of manufacturers of networked devices. For example, for Windows XP Professional network adapter is called: \ Device \ NPF_ (SE1DB_B4F6-D8F6BDC0E7B8) (Fig. 1). It is used because of special features determine the manufacturer's name for the network device and otheres (Fig. 2 shows the definition of a network adapter for Linux - the names 'any' and 'eth2' are due to the use of virtual machines from which Ubuntu 9.04 is runing).

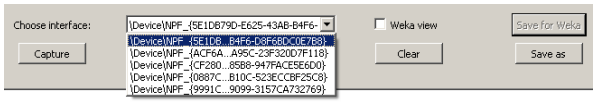


Figure 1. Available Windows Interfaces

Figure 2. Available Ubuntu interfaces (under virtual machine)

The library also allows to take detailed filtering of incoming packets (frames) of data at the hardware level (in this case, at the level of a network adapter). As you know, every time you receive data frames in its buffer, the network device generates a processor interrupt. Filtering is necessary: a network analyzer without the filter in the network of high bandwidth (above 10 Mbit/sec), with the graphical output make impossible work of other applications.

Pcap library provides a very simple interface to the BPF-filtration, concentrated in two main functions:

```
int pcap_compile(pcap_t *p, struct
bpf_program *fp, char *str, int
optimize, bpf_u_int32 netmask);
int pcap_setfilter(pcap_t *p, struct
bpf_program *fp).
```

The first function "compile" a string char * str, consisting of a set of regular expressions to filter in BPF-pseudocode and stores it in the list called struct bpf_program * fp, applying a filter to the network mask bpf_u_int32 netmask and "squeezing" it, discarding the redundant data in case of positive values of integer optimization. The second function uses the resulting pseudocode to the special device descriptor. Memory under bpf_program * allocates automatically.

There are many ways to read data packets from the open (and, optionally, filtrated) interface. The most suitable from the standpoint of saving resources is a function handler, caused when the new package (or group of packages) comes.

Function int pcap_loop (pcap_t * p, int cnt, pcap_handler callback, u_char * user). pcap_t * p - a handle to the open interface. int cnt - number of frames of data, that must be written before you pass control further.

Implementation of interfaces initialization is as follows:

```
while(!enough) {
rc =
pcap_dispatch(iface,number,packet_han
dler,NULL);
if(rc == -1) {
printf("pcap_dispatch() failed!\n");
return 1;   }}
```

In order to organize the analysis of network packets used to split the functions of these packages in two types - to provide the most complete information of the packet (IP, port numbers, details of the package itself, MAC address, country, etc.) and its output in a normal mode for easy viewing for user (Fig. 3), and the compatibility mode with the Weka tool (Fig. 4).

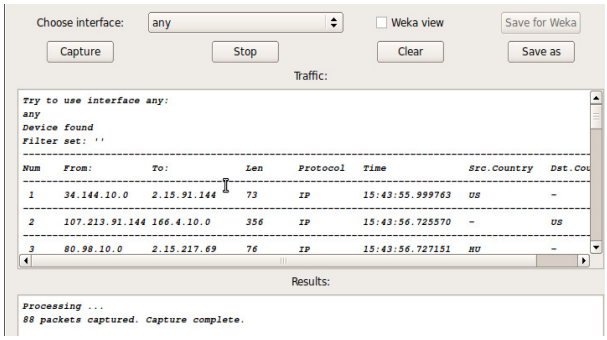


Figure 3. Capturing and parsing packets in Linux

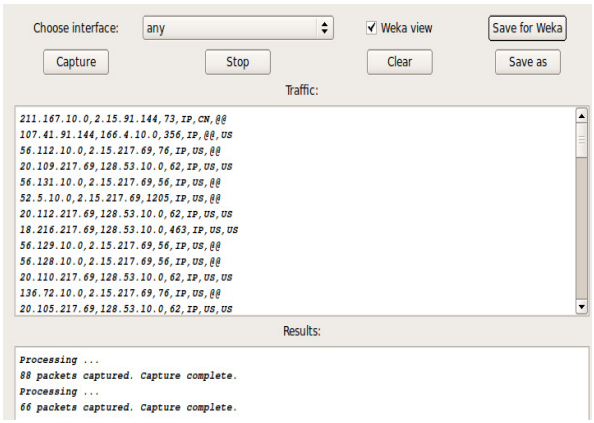


Figure 4. Capturing and parsing packets in Linux in WEKA compatible mode

3. WEKA

Weka (© The University of Waikato) is a collection of visualization tools and algorithms for data analysis and problem solving prediction, together with a graphical user shell access to them [3] [4]. Weka can perform such tasks feature selection, clustering (Fig. 6), classification, regression analysis and visualization of results.

Clustering of data packets can be used to separate groups of elements that can be potentially dangerous or

harmful, so using of this tool in the future will help developers of intrusion detection systems [5] [6] [7].

The algorithms of Weka are impossible without preliminary conversion of files to the special file format. This format is - Arff, which is presented in Figure 5. The header file provides a summary of information about the elements that follow further, since for the formation of these data is necessary to know all the session data to capture packets (Fig. 5).

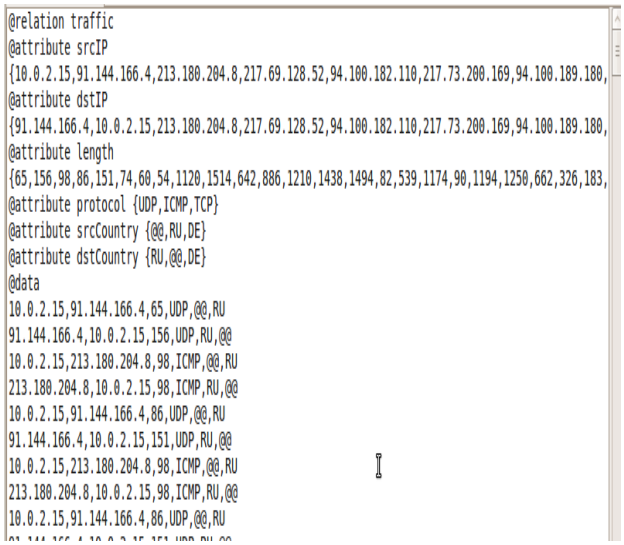


Fig. 5. Weka file format

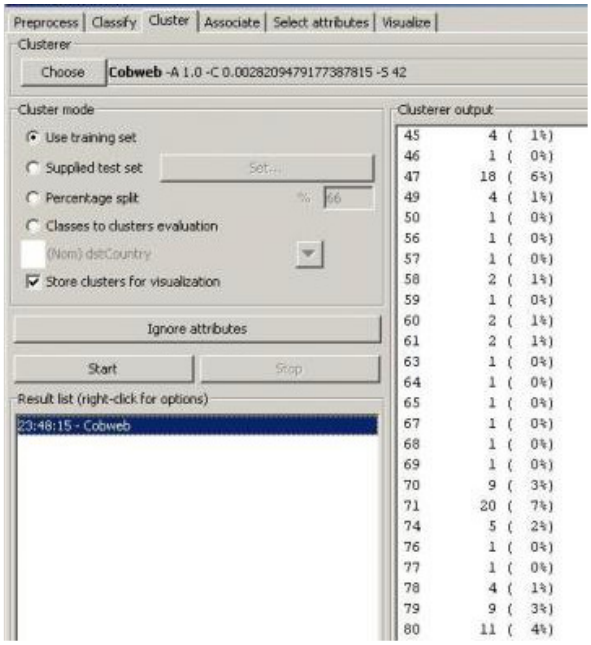


Fig. 6. Clustering packets using COBWEB algorithm

4. Conclusion

Using network traffic analyzer with artificial intelligence algorithms, the development of machine learning algorithms for these systems will improve security of corporate, finance and home networks.

References

1. W. Richard Stevens, Bill Fenner, Andrew M. Rudoff, UNIX Network Programming: The sockets networking API, Addison-Wesley Professional, 2004
2. Angela Orebaugh, Greg Morris, Ed Warnicke, Gilbert Ramírez, Ethereal packet sniffing, Syngress Media, Inc, 2004
3. Geoffrey I. Webb, Xing Huo Yui, AI 2004: advances in artificial intelligence, Springer, 2004
4. Douglas H. Fisher, Knowledge acquisition via incremental conceptual clustering, Springer Netherlands, 2004
5. Istvan Jonyer, Graph-based hierarchical conceptual clustering, The Journal of Machine Learning Research, 2002
6. Matrosov, AA, Architecture, network intrusion detection systems, IT security, 2005
7. Agranovsky A., Hadi RA, Balakin AV, teaching systems to detect and protect against intrusio

Development of software and mathematical model to control the regional system of fuel supply based on renewable energy sources

K.V. Ketova
Izhevsk State Technical University
Izhevsk, Russia
e-mail: ketova_k@mail.ru

A.S. Sayranov
Izhevsk State Technical University
Izhevsk, Russia
e-mail: shindou@mail.ru

E.V. Trushkova
Izhevsk State Technical University
Izhevsk, Russia
e-mail: chatongatin@mail.ru

R.Yu. Krivenkov
Izhevsk State Technical University
Izhevsk, Russia
e-mail: qwerty19106@yandex.ru

Abstract

The paper contains the developed scheme and proposes the algorithm for solving the problem of fuel supply for the distributed system of the region heat supply based on renewable fuel resources. The dynamic logistic task is solved based on the regional heat supply system of the Udmurt Republic.

1. Introduction

In recent years much attention has been paid to searching the alternative energy sources and developing new effective technologies for obtaining fuel-energy resources (FER) from renewable raw materials.

The assessment of energy stores in the UR demonstrated that it is advisable to use recycled wood wastes as an alternative fuel. The energy potential of wood wastes is about 140,000 tons per year, thus completely covering the demand of boiler rooms in the UR, which operate on the coal, electrical energy, etc instead of natural gas (slightly over 80,000 tons per year) [1].

According to the Concept of long-term social-economic development of the Russian Federation, the fuel based on wood wastes will be cheapest in the perspective till 2020. Wood chips after a special technological treatment are the most perspective type of such fuel.

When shifting to local renewable FER, the task of optimal fuel supply for the regional distributed heat supply system becomes the most urgent.

As the initial data required to solve the task of the region fuel supply, we use the information on the locations of lumbering and wood processing factories, habitations with heat sources (boiler rooms) that are planned to be shifted from traditional fuel types (coal, boiler oil, electric energy) to alternative ones.

2. Logistic scheme

The logistic scheme of heat sources supply with chips comprises 4 levels (see Fig. 1). Lumbering and wood processing factories accumulate wood wastes to be transported to waste accumulation points (WAP), where they undergo the initial mechanical treatment. The initially treated wood wastes are transported to fuel preparation points (FPP), where the chips are sorted out, heat treated and packed. Finally, the chips are transported to the consumers, i.e. heat sources of the UR. Warehouses for raw material storage are available on all the levels of the logistic scheme.

Since the demand in FER is changing during the heating period, the logistic task considered is dynamic.

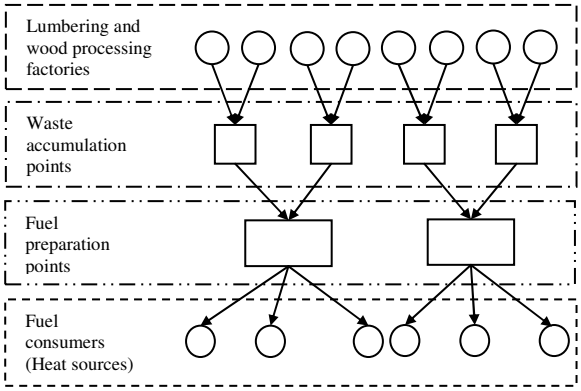


Figure 1. Scheme of regional fuel supply system

3. Logistic task

The algorithm for solving the dynamic logistic task of heat supply consists of five stages, which of them solves certain problems: routing, optimal distribution of resources at local and regional levels, store control at the warehouses of heat sources and points of FER preparation.

3.1. Routing task

Problem definition. The essence of solving the routing task consists in searching the optimal routes to transport wood wastes and FER. The minimization criteria are the transportation costs contained in the heat energy costs.

The following designations are introduced:

$G = \{g_i\}$ – aggregate of road intersection junctions, $i = \overline{1,m}$, m – number of junctions.

$S = \{s_{ij}\}$ – matrix of transportation costs for the delivery of one ton of raw material, where s_{ij} – costs for the delivery of raw material from i to j junction $i, j = \overline{1,m}$.

$$s_{ij} = c_{ij}^{l} d_{ij}, \quad (1)$$

where d_{ij} – distance from i to j junction; c_{ij}^{l} – specific tariff, cost of FER transportation along the l-type road. Specific tariff (rub/km) for the raw material transportation comprises depreciation costs of motor transport (a), fuel costs (u^l) and labor remuneration (w^l):

$$c^l = a + u^l + w^l, \quad (2)$$

at the same time $a = \frac{P}{L_r}$ (L_r – operational resource in km; P – price of the vehicle); $w^l = \frac{w}{v^l}$ (w – hourly labor remuneration tariff; v^l – movement speed along the road of l-type).

Let us define the route trajectory from the junction g_{k_1} to the junction g_{k_n} as $M_{k_1 k_n}^{k} = (g_{k_1}, g_{k_2}, ..., g_{k_n - 1}, g_{k_n})$, where g_{k_1} – initial junction of the route, g_{k_n} – final junction of the route, k_i – numbers of junctions in k route, $i = \overline{1,n}$.

Transportation costs for the delivery of raw material from the junction g_{k_1} to the junction g_{k_n} along the route k comprise the transportation costs for the delivery of raw material along each part of the road:

$$s_{k_1 k_n} = s_{k_1 k_2} + s_{k_2 k_3} + ... + s_{k_n - 1 k_n}. \quad (3)$$

The routing task consists in searching the route with minimal transportation costs for the delivery of 1 ton of raw material from the junction g_{k_1} to the junction g_{k_n}:

$$s_{k_1 k_n} \to \min. \quad (4)$$

Solution algorithm. To solve the routing task the complex 3-stage algorithm is proposed:

1. finding the initial route with the help of simplified Astar algorithm [2];
2. obtaining the multitude of allowed routes with the help of Yuen method for the arrangement of allowed routes;
3. finding the optimal route. The genetic algorithm using cross-breeding, selection and mutation operators adapted to the routing task is realized [3].

The application of crossing over operator is the most important part of the algorithm. The crossing over operator is applied if the common junction or common rib is available. Below is the algorithm of crossing over operator application for the common junction.

1. We assume that there are two different routes:

$$M_{k_1 k_n}^{1} = (g_{k_1}, g_{k_2}, ..., g_{k_n - 1}, g_{k_n}), \quad (5)$$

$$M_{l_1 l_n}^{2} = (g_{l_1}, g_{l_2}, ..., g_{l_n - 1}, g_{l_n}), \quad (6)$$

for which the following conditions are applicable:

$$g_{k_1} = g_{l_1} = g_n, \quad (7)$$

$$g_{k_n} = g_{l_n} = g_k, \quad (8)$$

$$g_{k_i} = g_{l_j} = g_h, \quad (9)$$

where g_n – initial junction of the routes in question; g_k – final junction of the routes in question; $g_{k_i} \in M_{k_1 k_n}^{1}$, $g_{l_j} \in M_{l_1 l_n}^{2}$ – intermediary junctions of the corresponding routes.

2. Crossing the routes (5), (6) by the junction g_h, we obtain the new routes:

$$M^{12} = (g_{k_1}, g_{k_2}, ..., g_{k_i - 1}, g_h, g_{l_j + 1}, ..., g_{l_n - 1}, g_{l_n}), \quad (10)$$

$$M^{21} = (g_{l_1}, g_{l_2}, ..., g_{l_j - 1}, g_h, g_{k_i + 1}, ..., g_{k_n - 1}, g_{k_n}). \quad (11)$$

3. We calculate the transportation costs of one ton of FER for the new routes. If the transportation costs for new routes (10), (11) are lower than for initial ones (5), (6), they participate in further solution of transportation problem.

Each stage of the algorithm possesses the acceptable operation speed and low complexity for computer implementation. In general, the algorithm allows avoiding the enumeration of all the possibilities of transportation routes.

3.2. Clustering task

Problem definition. The clustering problem is solved to define the location of WAP and FPP. The initial data to solve the clustering problem are the information on the locations of lumbering and wood processing factories, as well as habitations with heat sources that are planned to be shifted from traditional fuel types to alternative ones.

The essence of clustering analysis when defining the FPP locations consists in optimal combination of habitations with heat sources into the groups of closely located objects. In each cluster we should point out one habitation where FPP is planned to be placed under the condition of transportation costs minimization for FER delivery from FPP to the heat sources.

The following designations are introduced:
$H = \{h_i\}$ – multitude of habitations with heat sources, $i = \overline{1, M}$, M – number of habitations in question; $T^p = \{h_j^p\}$ – multitude of closely located habitations (p cluster), $p = \overline{1, K}$, K – number of clusters, h_j^p – j habitation in p cluster, $j = \overline{1, n_p}$, n_p – number of habitations in p cluster.

It is necessary to split the multitude $H = \{h_i\}$ into such groups $T^p = \{h_j^p\}$ to bring the total FER transportation costs to the minimum:

$$Z = \sum_{p=1}^{K} \sum_{j=1}^{n_p} s\left(h_l^p, h_j^p\right) \to \min, \quad (12)$$

where h_l^p – habitation where FRR is planned to be located; $s_{ij} = s(h_i, h_j)$ – FER transportation cost from i to j habitation.

In the same way the clustering task to define WAP location is set.

Solution algorithm. The general algorithm of cluster analysis comprises two stages: hierarchical cluster analysis is applied at the first stage, method of *k*-averages – at the second.

The hierarchical cluster analysis is applied to define the optimal number of clusters. The method of *k*-averages is applied for the distribution of objects by clusters and definition of optimal locations of WAP and FPP that are the centers of these clusters.

The heuristic Ardolan method is applied when defining the cluster center [4]. The methods takes into account the scattering of wastes produced and needs of heat sources.

3.3. Task of optimal distribution of resources in the area

Problem definition. The minimum of transportation costs for the delivery of wood wastes from WAPs to FPPs is the criterion of transportation problem minimization here:

$$Tr^{(1)} = \sum_{j=1}^{n} \sum_{i=1}^{m} c_{ij} d_{ij} q_{ij} \to \min, \quad (13)$$

where c_{ij} – specific tariff for the delivery of wastes from i WAP to j FPP; d_{ij} – distance from i WAP to j FPP; q_{ij} – wood waste volume taken from i WAP to j FPP; m – number of WAPs in the area; n – number of FPPs in the area.

The volume of wood waste deliveries is limited by the warehouse capacity of the corresponding points:

$$\sum_{i=1}^{m} q_{ij} \le Q_j^{FPP}, \; j = \overline{1, n}, \quad (14)$$

$$\sum_{j=1}^{n} q_{ij} \le Q_i^{WAP}, \; i = \overline{1, m}, \quad (15)$$

where Q_i^{WAP} – annual volume of wastes recycled at i WAP; Q_j^{FPP} – annual volume of waste consumption at j FPP.

3.4. Task of optimal distribution of resources in the region

Problem definition. Due to the difference in the volumes of FER production and consumption at the regional level, it can be necessary to solve the task of optimal distribution of excessive FER between FPPs of different areas.

The minimum of transportation costs for FER delivery between the areas is the criterion of transportation problem minimization in the region:

$$Tr^{(2)} = \sum_{j=1}^{N^-} \sum_{i=1}^{N^+} c_{ij} d_{ij} q_{ij} \to \min, \quad (16)$$

$$\sum_{i=1}^{N^+} q_{ij} \le Q_j^-, \; j = \overline{1, N^-}, \quad (17)$$

$$\sum_{j=1}^{N^-} q_{ij} \le Q_i^+, \; i = \overline{1, N^+}, \quad (18)$$

where N^+ – number of FPPs where the excessive FER is produced; N^- – number of FPPs lacking FER; q_{ij} – optimal volume of FER transportation from i FPP to j FPP lacking FER; Q_j^- – lack of FER at j FPP; Q_i^+ – excessive FER at i FPP.

3.5. Task of store control

The dynamic task of optimal store control consists in the task of control at the heat source warehouse and FPP warehouse.

Definition of the problem of optimal store control at the heat source warehouse. The essence of the solution of dynamic task of optimal store control at the heat source warehouse consists in searching optimal lot sizes and delivery time obtained as a result of minimization of the warehouse total costs.

The volume of fuel consumed by heat sources depends on the season, therefore we introduce the seasonality $s(t)$:

$$s(t) = \begin{cases} k_1, \; t \in [t_0; t_0 + \tau_1); \\ k_2, \; t \in [t_0 + \tau_1; t_0 + \tau_1 + \tau_2); \\ k_3, \; t \in [t_0 + \tau_1 + \tau_2; \, t_0 + \tau_1 + \tau_2 + \tau_3]. \end{cases} \quad (19)$$

Here k_m – seasonality factor; $m = 1, 2, 3$ – corresponds to autumn, winter and spring periods; τ_1, τ_2, τ_3 – duration of autumn, winter and spring

periods, respectively; t_0 – beginning of heating period; $t_k = t_0 + \tau_1 + \tau_2 + \tau_3$ – end of heating period, at the same time, the following condition is taken:

$$t_k - t_0 = \sum_{m=1}^{3} \tau_m . \quad (20)$$

The function $s(t)$ is normalized:

$$\int_{t_0}^{t_k} s(t)dt = 1 . \quad (21)$$

Then the season volume of FER consumption for i heat source will be defined by the formula:

$$d_{im} = d_i k_m , \quad (22)$$

where d_i – FER consumption volume by i boiler-room for a year.

To define the general seasonal costs comprising the fuel purchasing, delivery and storage, the following formula is applied [5]:

$$C_{im}(q_{im}, t) = c_{im} d_{im} + \frac{z_{im} d_{im}}{q_{im}} + \\ + h_{im}\left(\frac{q_{im}}{2} + q_{im}^{p}\right), \quad (23)$$

where c_{im} – FER specific cost for i boiler-room in m season; z_{im} – organizational costs for a lot for i boiler-room in m season; h_{im} – specific costs for FER storage at the warehouse for i boiler-room in m season; q_{im} – size of FER lot delivered for i boiler-room in m season; q_{im}^{p} – FER reserve at the warehouse for i boiler-room in m season.

The solution of the problem of optimal store control for i heat source in m season consists in selecting the parameter q_{im} to minimize the total costs:

$$C_{im}(q_{im}, t) \to \min . \quad (24)$$

We will define the optimal solution of this problem as $q_{im} = q_{im}^{*}$. At the same time, the duration of the optimal cycle of store renewal is defined by the formula:

$$\tau_{im}^{*} = \tau_m \frac{q_{im}^{*}}{d_{im}} . \quad (25)$$

Definition of the problem of optimal store control at FPP. The essence of the solution of dynamic task of optimal store control at FER preparation point consists in searching optimal lot sizes and delivery time obtained as a result of minimization of the warehouse total costs.

The demand N of heat sources in FER served by j FPP in m season equals:

$$D_{jm} = \sum_{i=1}^{N} d_{im} . \quad (26)$$

FER productivity at j FPP in m season is designated as P_{jm}, then FER store at it increases at the speed $P_{jm} - D_{jm}$. Therefore the maximum store in the end of the period is calculated by the formula:

$$M_{jm} = (P_{jm} - D_{jm})\tau_m . \quad (27)$$

At the same time, it is necessary to transport Q_{jm} tons of wood wastes from WAP to FPP:

$$Q_{jm} = P_{jm}\tau_m . \quad (28)$$

From (27), (28)

$$M_{jm} = \frac{(P_{jm} - D_{jm})Q_{jm}}{P_{jm}} . \quad (29)$$

Total seasonal costs at j FPP comprising wood waste purchasing, delivery and storage are found by the formula [5]:

$$C_{jm}(Q_{jm}, t) = c_{jm} D_{jm} + \frac{z_{jm} D_{jm}}{Q_{jm}} + \\ + h_{jm} \frac{(P_{jm} - D_{jm})Q_{jm}}{2P_{jm}} , \quad (30)$$

where c_{jm} – specific cost of wood wastes for j FPP in m season; z_{jm} – organizational costs for a lot for j FPP in m season; h_{jm} – specific costs for wood waste storage at FPP warehouse for j FPP in m season.

The solution of the problem of optimal store control at the warehouse of j FER preparation point in m season consists in searching the optimal lot size Q_{jm}^{*} and delivery time τ_{jm}^{*} obtained as a result of minimization of the warehouse total costs.

4. Conclusion

Thus the algorithm for solving the dynamic problem of fuel supply logistics consisting of five interconnected logistic tasks was developed. The algorithm was realized and used as the basis of information-analytical system "Fuel supply in the UR".

The information-analytical system is based on the data manager MS SQL Server 2008. The electronic map of the UR was developed in the medium GIS MapInfo 9.5 for the project graphic and visual support. The infrastructure objects, such as habitations, heat sources, lumbering and wood processing factories, cattle breeding farms, transport network were applied to the topological base of the UR. The program complex database contains the information on the main characteristics of these objects.

Thus the fuel supply scheme of one of the UR areas is demonstrated in Fig. 2.

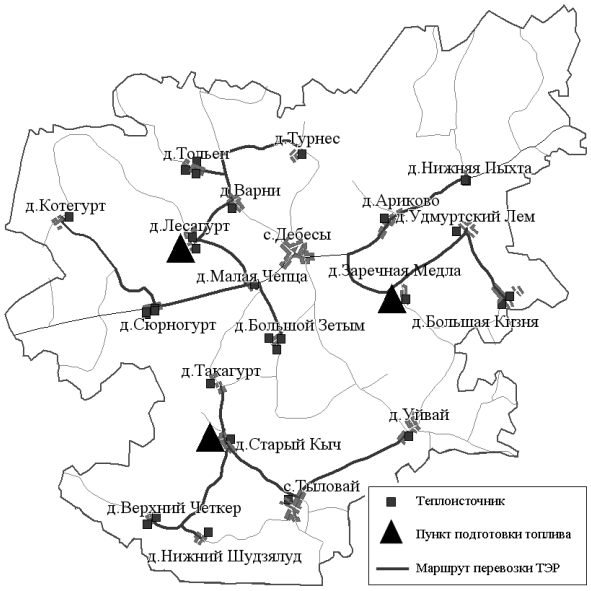


Figure 2. Area fuel supply scheme

All the boiler-rooms in the area that will undergo the reconstruction are divided into three groups by the method of cluster analysis. In each cluster the habitation where the FPP will be located is defined by Ardolan method. Similarly the number and locations of wood waste accumulation points are defined based on the information on the deforestation zones in the area. The optimal routes of FER transportation from fuel preparation points to heat sources are defined while solving the routing problem.

As a result of solving the task of the logistics of heat supply regional system, the optimal routes of wood waste and FER transportation, the locations of waste accumulation points and fuel preparation points, as well as optimal schedules of raw material delivery to the warehouses of FPP and heat sources in the Udmurt Republic are defined.

References

1. Savelyev V.A., Presnukhin V.K., Rusyak I.G. Concept of fuel-energy complex development of the Udmurt Republic in 2003-2010. Izhevsk: ISTU, 2002. – 260 p.
2. Kristofides N. Graph theory. Algorithmic approach. M.: Mir, 1978. – 432 p.
3. Gladkov L.A., Kureychik V.M. Genetic algorithms. M: Physmathlit, 2006. – 320 p.
4. Prosvetov G.I. Mathematical methods in logistics. M.: Alfa-Press, 2008.– 230 p.
5. Shikin E.V., Chkhartishvili A.G. Mathematical methods and models in management. M.: Delo, 2004.

Development of dataware-computational program for the diagnostics and treatment of cardiovascular system diseases based on mathematical and computer 3D modeling

Sufiyanov V.G.
Faculty of Applied Mathematics
Izhevsk State Technical University
Izhevsk, Russia
e-mail: vsufiy@gmail.com

Igosheva K.P.
Faculty of Applied Mathematics
Izhevsk State Technical University
Izhevsk, Russia
e-mail: primat@istu.ru

Vasilyeva I.V.
Faculty of Applied Mathematics
Izhevsk State Technical University
Izhevsk, Russia
e-mail: primat@istu.ru

Abstract

The problems of mathematical and computer 3D modeling of cardiovascular system is discussed. The software for mathematical 3D modeling of biological systems is reviewed. The mathematical Aliev-Panfilov model of the excitation electrical wave spread in virtual 3D anatomical model of heart is implemented. The statistical models for determining the coronary arteries according to stress echocardiography are prepared. The program for the input of data on the affected parts of the heart and display of the predicted lesions of coronary arteries in 3D model of the heart is developed.

1. Introduction

Cardiovascular system – one of the major biological systems, which has many functions and is linked with the work of other organs and systems, and in turn, bears a heavy burden related to human activity. Despite numerous protective mechanisms the cardiovascular system is exposed to such diseases as: heart attacks and strokes, hypertension, arrhythmia, etc. The mechanisms that lead to such disorders in some cases are known, but still many unsolved problems connected with the emergence and course of diseases remain. This is due to the fact that heart and cardiovascular system are an open system connected with endocrine, nervous, musculoskeletal and other human systems. The disorder in cardiovascular system leads to the medical condition of these systems and vice versa.

In this regard it is necessary to solve complex problems not only of treatment but also prevention of the diseases of cardiovascular system. To do this we need to know the mechanisms of these diseases, properly diagnose and identify their causes. To address these challenges doctors use empirical methods - conducting mass surveys, modeling *in vivo.* The implementation of these studies is required to obtain reliable information about the processes taking place at the methodological level, to compare these data with theoretical knowledge. However, when establishing the diagnosis, a doctor has to deal not with a set of observations, but with a specific patient and the need to properly diagnose the case and determine the cause of disease. Therefore, for the correct diagnosis and determination of the causes of illness a doctor should be allowed to simulate the development of the disease *in silico* on "a virtual patient", based on mathematical modeling on a computer with the given set of parameters and characteristics of the person, the data of his medical history, clinical and diagnostic data, laboratory and instrumental studies. Our goal is to develop a multifunctional computer system for the simulation of cardiovascular system in general.

Software systems in cardiovascular system simulation were surveyed. One of the major projects in the field of mathematical modeling of biological systems is the international Physiome Project [1], which combines the institutions and universities around the world. Mathematical models of cardiovascular system have been implemented INRIA [2], for example, CardioSense3D [3] allows visualizing the results of magnetic resonance imaging of the heart.

Despite the large range of developed software systems for mathematical modeling of biological systems, in our study we used a more universal program for solving problems. It is connected with the fact that in the software systems developed either simple processes can be simulated or it is impossible to improve the functional characteristics of the developed models.

2. Electrical model of heart

Conducting system is the primary source of other systems, since it is due to propagating electrical impulses that other functions of the heart are carried out: mechanical compression and blood movement through arteries.

For the first time, the ionic model of the distribution of the excitation pulse appeared in the papers of

biophysicists Hodgkin A.L. and Huxley A.F. [4]. The rheological models of electric pulse distribution are based on Hodgkin-Huxley mode [5]: FitzHugh-Nagumo, Aliev-Panfilov, Zeeman and Biktashev. We considered Aliev-Panfilov monodomain mathematical model of excitation electrical wave distribution [5]:

$$\begin{cases} \dfrac{\partial u}{\partial t} = f(u,v) + \Delta u, \quad in \quad \Omega \times [0,T]; \\ \dfrac{\partial v}{\partial t} = g(u,v), \qquad in \quad \Omega \times [0,T]; \end{cases} \tag{1}$$

where u – normalized dimensionless value corresponding to transmembrane potential; v – dimensionless value corresponding to slow current recovery; t – time; Δ – Laplace operator; Ω – integration domain; T – integration time. Functions $f(u,v)$ and $g(u,v)$ in Aliev-Panfilov model have the form:

$$\begin{cases} f(u,v) = -ku(u-a)(u-1) - uv; \\ g(u,v) = -\left(\varepsilon_0 + \dfrac{\mu_1 v}{\mu_2 + u}\right)(v + ku(u-a-1)). \end{cases} \tag{2}$$

The model used the parameters corresponding to the heart muscle $k = 8.0$, $\varepsilon_0 = 0.01$, $a = 0.15$, $\mu_1 = 0.2$, $\mu_2 = 0.3$ [5]. The wave of excitation is induced in atrioventricular bundle and extends to the left and right ventricles (Fig. 1).

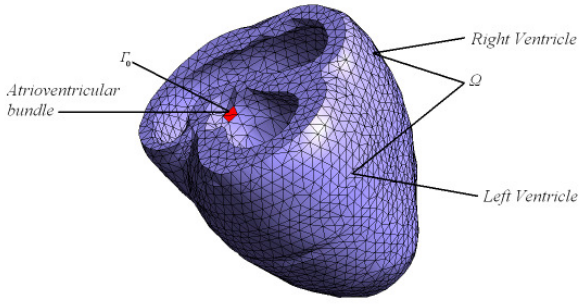


Fig. 1. Triangulation of the left and right ventricles of the heart

Equation (1) is supplemented by initial conditions:

$$\begin{cases} u(0) = 0, \quad in \quad \Omega; \\ u(0) = 1, \quad in \quad \Gamma_0; \\ v(0) = 0, \quad in \quad \Omega; \end{cases} \tag{3}$$

and boundary conditions:

$$\begin{cases} \delta \cdot u + \dfrac{\partial u}{\partial \vec{n}} = \delta, \ in \ \Gamma_0 \times [0,T]; \\ \dfrac{\partial u}{\partial \vec{n}} = 0, \qquad in \ \Gamma_1 \times [0,T]; \\ \dfrac{\partial v}{\partial \vec{n}} = 0, \qquad in \ \Gamma \times [0,T]. \end{cases} \tag{4}$$

where $\vec{n}$ – normal vector to the surface; t_1 – time of the pulse; $\delta = \delta(t < t_1)$ – function equal to 1 at $t < t_1$ and to 0 at $t \geq t_1$; $\Gamma = \Gamma_0 \cup \Gamma_1$ – Ω surface area, consisting of the area of impulse occurrence Γ_0 and Γ_1 – other surfaces except for Γ_0. In the example, the time of the pulse t_1 considered as equal to 3 time units. To solve the system of equations (1)-(4), we represent it in weak formulation:

$$\begin{cases} \int\limits_\Omega \dfrac{\partial u}{\partial t} \cdot u' \, dV + \int\limits_\Omega \nabla u \cdot \nabla u' \, dV + \int\limits_{\Gamma_0} \delta u \cdot u' \, dS = \\ = \int\limits_\Omega f(u,v) \cdot u' \, dV + \int\limits_{\Gamma_0} \delta \cdot u' \, dS; \\ \int\limits_\Omega \dfrac{\partial v}{\partial t} \cdot v' \, dV = \int\limits_\Omega g(u,v) \cdot v' \cdot dV, \end{cases} \tag{5}$$

where u' and v' are testing functions.
The area triangulation (Fig. 1) was in the medium Gmsh [6]. Aliev-Panfilov model was implemented in the medium GetDP [7]. The results of the model implementation from 0 ms to 60 ms are shown in Fig. 2.

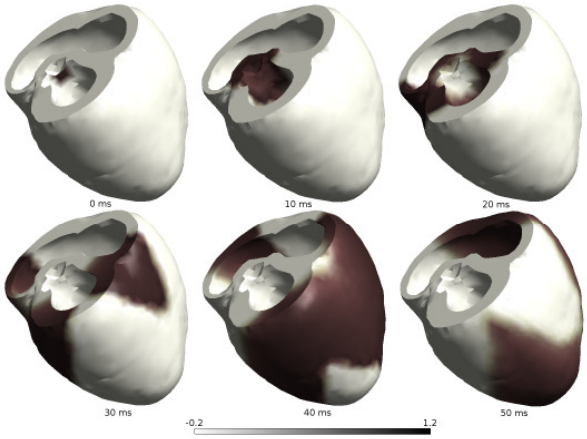


Fig. 2. Distribution of transmembrane potential at time intervals 0, 10, 30, 40 and 50 ms

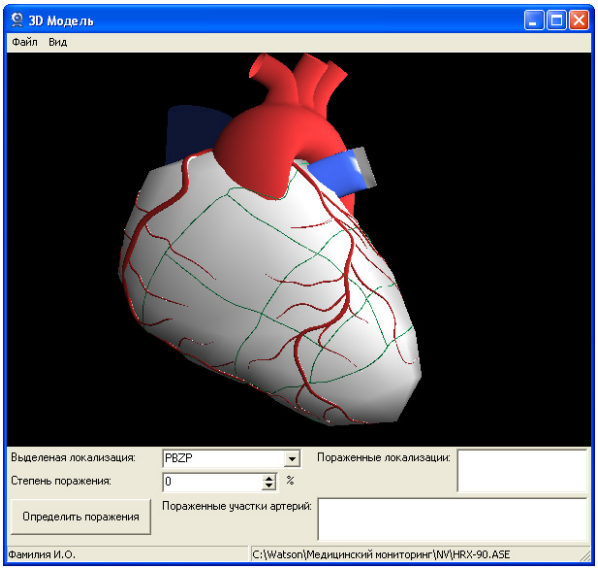
The results shown in Fig. 2 give only a superficial understanding of excitation wave propagation. The next step is to refine it adding the conducting system to Aliev-Panfilov model. Afterwards "the virtual electrocardiogram" should be simulated. This will create a set of electrocardiograms for people with

different body types and heart topology, different ages, in normal health and with medical conditions.

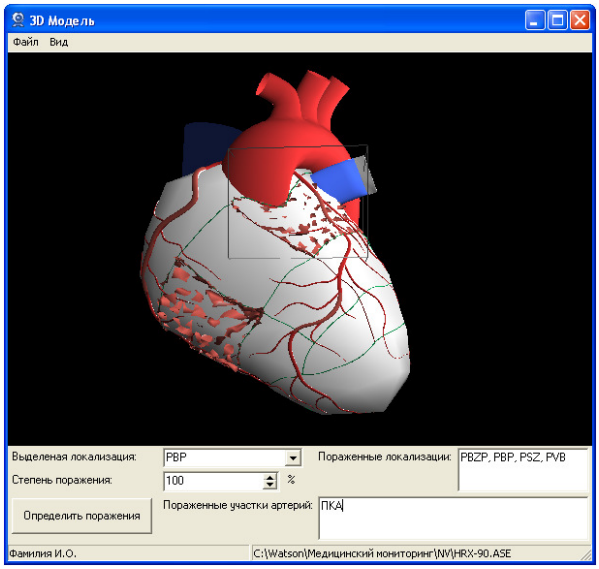
3. Statistical model to diagnose the coronary artery atherosclerosis

On the basis of heart electrical model it is planned to create a hydraulic model, in which the heart will act as a pump, and arteries and veins – the hydraulic system. The heart work is supported by the blood supply system through coronary arteries. If you have problems, some areas of the heart muscle stop contracting in the appropriate tone. Basically, this is caused by atherosclerosis. The investigation of the relationship of mechanical properties of myocardium and hydraulic characteristics of the heart work will solve several problems associated with the diagnosis of coronary artery atherosclerosis by the data of stress echocardiography.

Together with the doctors of the Republican clinical diagnostic center (Izhevsk) the studies were conducted related to the identification of areas of constrictive affections of coronary arteries according to stress echocardiography data. In the course of studies the classification models based on logistic regression [8], neural networks, logic and algebraic methods and decision trees were constructed. The program for the visual input of the data on heart section hypokinesis was designed, and these data are used to obtain the information about the affected arteries based on the developed statistical decision rules (Fig. 3).



a) without hypokinesis



b) hypokinesis

Figure 3. 3D heart model

4. Perspectives and future development

Diseases of cardiovascular system are both the cause and consequence of the diseases of other organs and systems, since the human body is a developing, self-adjusting adaptive system, which involves the nervous and endocrine systems. The role of information sources play physical, chemical and biological characteristics of the organs and systems. Based on 3D models developed and literature and experimental data available, it is planned to develop a system-wide mathematical model for the regulation and adaptation of the human body. This will solve several problems associated with the modeling of dynamically developing diseases, such as hypertension, myocardial infarction, ciliary arrhythmia, as well as address issues related to their prevention, diagnosis, treatment and development prognosis.

5. Conclusion

Thus, the virtual 3D model of the heart was developed and the excitation wave propagation in this model was simulated. The results presented in this paper give only a superficial understanding of excitation wave propagation in the heart, as they do not take into account the conducting system. The inclusion of this system into the model will produce more realistic results and will allow using it when solving a number of medical problems.

The investigation of the relationship between mechanical characteristics of the myocardium and coronary blood flow hydraulic characteristics allowed developing logic models to identify the affection of coronary arteries according to stress echocardiography data. These decision rules have been framed as a

program that implements a 3D interface for entering data on the affected areas of the heart and displaying diagnostic results.

References

1. Physiome Project. – URL: http://www.physiome.org/
2. Boulakia M., Cazeau S., Fernandez M.A., Gerbeau J.-F., Zemzemi N. Mathematical Modeling of Electrocardiograms: A Numerical Study. – Institute National De Recherche en Informatique et en Automatique. – Raport de recherche № 6977. - 2009. – 42 p. – URL: http://www-sop.inria.fr/CardioSense3D/spip.php?article163
3. Cardiosense3D // Institute National De Recherche en Informatique et en Automatique: [сайт]. – URL: http://www-sop.inria.fr/CardioSense 3D/
4. Hodgkin A.L., Huxley A.F. A quantitative description of membrane current and its application to conduction and excitation in nerve. Journal of Physiology, 117:500–544, 1952.
5. *Elkin Yu.E.* The autowave processes // Mathematical biology and informatics. – 2006. – V. 1. – № 1. – pp. 27-40.
6. Geuzaine C., Remacle J.-F. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities // International Journal for Numerical Methods in Engineering, Vol. 79, Issue 11, pp. 1309-1331, 2009. – URL:http://geuz.org/gmsh/gmsh_paper_preprint.pdf
7. Dular P., Geuzaine C., Henrotte F., Legros W. A general environment for the treatment of discrete problems and its application to the finite element method // IEEE Transactions on Magnetics, Vol. 34, Num.5, pp. 3395-3398, 1998.
8. Multanovsky B.L., Petrov A.G., Sufiyanov V.G., Korolev, D.A. Method for determining the localization of hemodynamically significant stenosis of the heart arteries. – Patent of the Russian Federation for the invention No 2352255 dated 20.04.2009.

The Corporate Information Systems Evolution

Blagodatsky Gregory
ISTU
Udmurt Republic, Russian Federation
e-mail: blagodatsky@gmail.com

Abstract[1]

Paper gives short review of corporate management system evolution from Grant's diagram to E-business [5, 6, 14, 18].

1. Introduction

Most business leaders are trying to find a universal solution that can immediately eliminate the whole set of problems. Each year there is a new concept, the new method. However, there is one "but": there is no universal solutions. Everything depends on ourselves, and the real way is improvement of existing enterprise management approaches.

2. Paper text

The classical MRP II / ERP system involves relationship of planning and management functions with the functions of planning, accounting and orders, suppliers, production, customer, financial management. In the CIM system computer aided design systems (CAD systems) and process control system (PCS system) to all this are added. While earlier such a close interaction with the business system is not expected for these functions [15, 19]. Fig. 1 shows the general scheme of CIM.

Summary up, it is possible that CIM system integrates various software products, that have different ideology, operating systems and data formats. The very problem is the most difficult, but solvable. But, in practice, simple integration of different programs is not enough, because it must be cooperation and collaboration of enterprise divisions, that often conflict with each other.

To a large extent, the development of this technology was contributed by the establishment of international standards, which only recognizes some of the data exchange formats, and the creation of special programs - converters to facilitating information exchange. It is believed that in the future CIM will cover more applications developed in accordance with its standards or adapt to them. This process will continue so long as full and unrestricted interaction of any program from any manufacturers will be achieved.

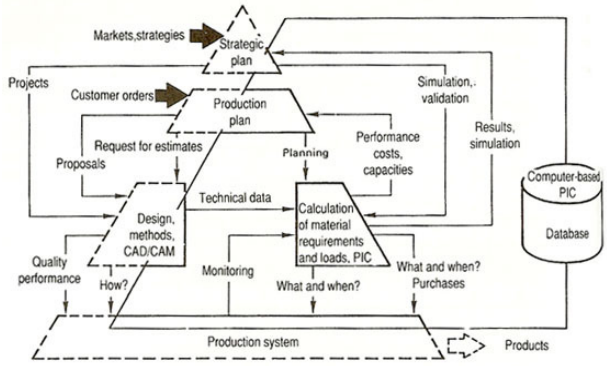


Figure 1. CIM general scheme

It is significant to say that the introduction of these systems are at a very slow. The reason is - the high cost of such decisions.

There are two management concepts in this area to date:

- Electronic business – ability to interacts with trading partners through Internet [14].
- CRM (Customer Relationship Management) is the management of interaction with customers [10].

These terms taken both individually and together sound quite attractive. The vendors of these systems say that they greatly increase the competitiveness of companies. However, they forget to say or deliberately silent that the possibility of quickly receiving orders through the Internet and information of the entire history of the relationship with the clients are not enough to deter them, or take a new one. There is a need for more reliable base, that including effective management system (JIT, MRP, MRP II, ERP, APS and CIM). It allows to reduce the percentage of damage production, expedite fulfillment, improve quality, lower production costs. Using such system, the company can serve customers faster, while spending less resources and sell products at a lower price.

3. Conclusion

Finally most business leaders are trying to find a universal solution that can immediately eliminate the whole set of problems. Each year there is a new concept, the new method. However, there is one "but":

[1] **Proceedings of the International Workshop "Innovation Information Technology - Theory and Practice", September 6th-10th, Dresden, Germany, 2010**

there is no universal solutions. Everything depends on ourselves, and the real way is improvement of existing enterprise management approaches.

This may contribute to specific tools, managing information. These tools are:

MRP - the method of calculating necessary for the implementation of main production plan material's volume and component's quantity [2].

MRP II - the method of calculating required execution resources to perform an MPP (materials and components, production facilities and the ability of personnel, finance, etc.). The concept of management and class of enterprise information management systems that manage the main areas of company activity such as planning, production, finance, procurement and sales [8].

JIT - management philosophy, which is based on the principle: continually looking for processes that increase cost, while not increase the consumer cost of products, and eliminate them [3, 12, 16, 17].

TOC - the theory of constraints. There is the concept of management, administering "bottlenecks" [7, 8].

APS - a management method of new level. It allows quick intervening in the production process by making the necessary adjustments in the production plan of arbitrarily often. in time, it can displace MRP II [4, 11].

CRM - the management method, in account of the individual requirements of every customer to product. This technology allows to personalize products in detail [10].

ERP - enterprise management information system, that is based on MRP II management algorithms, implementing JIT and TOC methodologies. Modern ERP systems combination of APS and CRM modules and e-business solutions.

Announced the death of the above methods for over 10 years, but they are continue to be used. Indeed, today there aren't something new. I'm suppose that for Russian companies, their use is just in time now.

References

1. Advanced Planning and Scheduling Solutions in Process Industry / Edit. Hans-Otto Ganther, Paul van Beek. – Cambridge: Springer, 2003. – 426 p.
2. APICS dictionary / Edit. J. F. Cox, etc. – Chicago: American Production and Inventory Control Society, 1992. – 144 p.
3. Ford, Henry. My life and work. / Henry Ford. – Sioux Falls: NuVision Publications, 2007. – 192 p.
4. Gaither, Norman. Production and operations management / Norman Gaither, Gregory V. Frazier. – Cincinnati: South-Western College Publishing, 1999. – 846 p.
5. Gantt, Henry L. A graphical daily balance in manufacture. / Henry L. Gantt. // Transactions of the American Society of Mechanical Engineers. – 1903. – Volume XXIV, – P. 1322-1336.
6. Gantt, Henry L. Work, Wages, and Profits / Henry L. Gantt. – New York: Engineering Magazine Co., 1916. – 330 p.
7. Goldratt, Eliyahu M. The Goal: A Process of Ongoing Improvement / Eliyahu M. Goldratt, Jef Cox. – Great Barrington: North River Press, 2004. – 384 p.
8. Goldratt, Eliyahu M. The Race / Eliyahu M. Goldratt, Robert E. Fox. – Aldershot: Gower Publishing Ltd., 1994. – 181 p.
9. Keller, Erik L. Enterprise Resource Planning : The changing application model / Erik L. Keller // SAP R/3 3.1. – Stamrord: Gartner Group, 1996. – P.8.
10. Kincaid, Judith W. Customer Relationship Management : Getting It Right / Judith W. Kincaid. – New York: HP Books and Prentice Hall PTR, 2002. – 512 p.
11. Korem, Yoram. Computer Control of Manufacturing Systems / Yoram Korem. – New York: McGraw Hill Inc., 1983, – 304 p.
12. Liker, Jeffrey. The Toyota Way : 14 Management Principles From The World's Greatest Manufacturer / Jeffrey Liker. – New York: McGraw-Hill Inc., 2003. – 350 p.
13. Orlicky, J. Material Requirements Planning : The New Way of Life in Production and Inventory Management / J. Orlicky. – New York: McGraw-Hill Inc., 1975. – 292 p.
14. Papazoglou, Mike. e-Business : Organizational And Technical Foundations / Mike Papazoglou, Pieter Ribbers. – Indianapolis: John Wiley & Sons, 2006. – 750 p.
15. Rehg, James A. Computer-Integrated Manufacturing / James A. Rehg , Henry W. Kraebber. – New Jersey: Prentice Hall, 2004. – 592 p.
16. Taiichi, Ohno. Just-In-Time for Today and Tomorrow / Ohno Taiichi. – New York: Productivity Press, 1988. – 145 p.
17. Taiichi, Ohno. Toyota Production System : Beyond Large-Scale Production / Ohno Taiichi. – New York: Productivity Press, 1988. – 163 p.
18. Taylor, Frederick Winslow. The Principles of Scientific Management / Frederick Winslow Taylor. – New York: Dover Publications Inc., 1998. – 76 p.
19. Waldner, J. CIM : Principles of Computer Integrated Manufacturing / J. Waldner. – Indianapolis: J. Willey & Sons, 1992. – 206 p.