VARNA FREE UNIVERSITY "CHERNORIZETS HRABAR"
INSTITUTE OF MATHEMATICS AND INFORMATICS OF THE
BULGARIAN ACADEMY OF SCIENCE

INFORMATICS
IN THE SCIENTIFIC KNOWLEDGE

INTERNATIONAL SCIENTIFIC CONFERENCE

2012

From 27 to 29 June 2012 the Fourth International Conference "Informatics in Scientific Knowledge" - ISK'2012, organized by Varna Free University "Chernorizets Hrabar" and the Institute of Mathematics and Informatics of the Bulgarian Academy of Science was carried out in Varna. The achievements in application of informatics and information technologies in scientific knowledge were discussed in the presented papers. Special attention will be paid to the information support in teaching and research work.

Main Topics:
- Informatics in the Economics, Management and Culture
- Informatics in the Public Administration, Law and Corporative Security
- Informatics in Engineering and Technologies
- Informatics in Education
- Mathematical Basics of Informatics

The conference waschaired jointly with
Prof. Dr. Sc. Anna Nedyalkova
Rector of Varna Free University "Chernorizets Hrabar"
and
Acad. Prof. Dr. Sc. Stefan Dodunekov
Head of the Bulgarian Academy of Sciences

INTERNATIONAL PROGRAMME COMMITTEE

- **Chairs:**

Stefan Dodunekov	Bulgaria
Nikolay Lyutov	Bulgaria

- **Members:**

Alberto Arteta	Spain
Alexey B.Petrovsky	Russia
Ali Güngör	Turkey
Angel Castellanos	Spain
Cvetko Andreeski	FYR Macedonia
Dana Petcu	Romania
Daniel Jolly	France
Evgeniya Cheremisina	Russia
Evgeny V.Pashkov	Ukraine
Hasmik Sahakyan	Armenia
Hassane Abuaissa Artoa	France
Iryna Pliss	Ukraine
Ivan Jelinek	Czech Republic
Jie Lu	Australia
Leon Rothkrantz	Netherlands
Levon Aslanyan	Armenia
Liudmila Cheremisinova	Belarus
Marcello Giacomantonio	Italy
Mile Stankovski	FYR Macedonia
Norvuz Allahverdi	Turkey
Nuria Gómez Blas	Spain
Oleg Senko	Russia
Panagiotis Bamidis	Greece
Piet Kommers	Netherlands
Ravil Muhamediev	Khazahstan
Tatjana Kolemishevska-Gugolovska	FYR Macedonia
Victors Gopejenko	Latvia
Vladimir Korenkov	Russia

Avram Eskenazi	Bulgaria
Dimitar Damyanov	Bulgaria
Dragomir Dobrudzhaliev	Bulgaria
Edi Chakarov	Bulgaria
Emil Nikolov	Bulgaria
Krasimisr Manev	Bulgaria
Krasimir Markov	Bulgaria
Krasimira Stoilova	Bulgaria
Lyubomir Sotirov	Bulgaria
Rady Romanski	Bulgaria
Radoslav Yoshinov	Bulgaria
Rusalin Nikolaev	Bulgaria
Teodora Bakardjieva	Bulgaria
Todor Stoilov	Bulgaria

TECHNICAL COMMITTEE

- **Chairs :**

Zdravko Slavov	Bulgaria
Krasimir Markov	Bulgaria

- **Scientific secretary:**

Maria Bruseva	Bulgaria
Neli Maneva	Bulgaria

- **Organization secretary:**

Antonina Ivanova	Bulgaria

- **Technical secretary :**

Pavleta Krasteva	Bulgaria

- **Members:**

Krasimir Trichkov	Bulgaria
Martin Tcenov	Bulgaria
Nikolaj Dimitrov	Bulgaria
Dimitar Ivanov	Bulgaria
Veselina Spasova	Bulgaria
Galina Momcheva	Bulgaria

CONTENT

INFORMATICS IN THE ECONOMICS, MANAGEMENT AND CULTURE

INFORMATICS IN THE PUBLIC ADMINISTRATION, LAW AND CORPORATIVE SECURITY

INFORMATICS IN ENGINEERING AND TECHNOLOGIES

INFORMATICS IN EDUCATION

MATHEMATICAL BASICS OF INFORMATICS

INFORMATICS IN THE ECONOMICS, MANAGEMENT AND CULTURE

USING SIMULATION FOR RISK ANALYSIS IN MS EXCEL

Mariya Bruseva

Abstract: Present work describes the main groups of methods for risk assessment and his qualitative and quantitative determination. It's created summary of simulation methods and a Monte Carlo simulation. With a specific example the author demonstrate modeling simulation using MS Excel functions.
Key words: Risk assessment, risk analysis, Monte Carlo simulation, MS Excel functions, Distribution

Introduction

Every enterprise is inextricably linked to the concept of 'risk'. Depending on the specifics of their business, market and political conditions, as well as the development strategy in its business enterprise is faced with different types of risk. The reason for the existence of risk can be economic crises, natural disasters, computer viruses and other phenomena and events that may lead to failure to achieve business purposes. However, risks can be managed as well as processes of production. For the successful existence of the enterprise, the developer must be committed to implement technical innovations to take bold, non-trivial action, which naturally increases the risk. It is therefore necessary to correctly assess the degree of risk and his management, which will lead to more efficient market outcomes.

1. Analysis and risk assessment.

Risk analysis – these are the procedures of identification of risk factors and assessing their significance, analysis of the probability of undesirable events which occur that affect the achievement of project objectives. Risk analysis includes risk assessment and respective methods for reducing the risk or reduction of the associated adverse effects. Risk assessment – this is a calculation with qualitative or quantitative methods degree of risk.

In the process of risk analysis is used effective modeling which approach allows studying dynamics of the external and internal environment of the enterprise, the replication of the actions of various

factors and conditions of the economic environment, to study the effectiveness of alternatives to the project, to identify possible errors and correcting them at this. In these actions should not cause material losses, not apply real material loss [1], [2], [3], [8]

General principles of risk analysis are: full coverage of the research field, reporting the strategy of the enterprise, reporting "time" factor, reliability of the received information, reduction of the methods used to detect risk

2. Methods of risk analysis

Risk analysis can be divided into two complementary types: qualitative and quantitative. Purpose of the qualitative is to identify causing risk areas factors and the types of risk. The final results of the qualitative risk analysis are used to input information to perform a quantitative analysis. Quantitative analysis should allow for numerical determination of the individual risks and risks of the enterprise as a whole. The implementation of quantitative risk assessment facing great difficulties related to the fact that this evaluation needs relevant starting information.

3. Quantitative determination of risk

In world practice of financial management are used different methods to quantifying risk. The most popular used are as follows:

3.1 Method of correction of discount rate

3.2 Method of credibility equivalents

3.3 Sensitivity analysis (the base is analyzing the effectiveness criteria - PP, NPV, IRR, IP)

3.4 Method of scenarios

3.5 Analysis of the probability distributions of the payment flows

3.6 Method "Tree of probability Decisions"

3.7 Monte Carlo simulation

4. Simulation

Simulation is one of the most powerful methods for analysis of economic systems. Overall, the simulation is the process of conducting computer experiments with mathematical models of complex systems in the real world. The objectives of these experiments can be various - determining the properties and regularities of the system, finding solution of the specific practical problems etc. As result of the development of computer technology and software, the use of simulation in the economy

has grown significantly. Currently, it is used for solving the management company problems and also for modeling at the macroeconomic level.

Imitation is a computer experiment. The only difference between this experiment and the reality is that this is done with a model of the system and not with the system itself. The conducting experiments with a real economic system is at least unwise, costly and practically difficult. Thus, the simulation is the only way to explore the system without a real experiment.

Often it is practically impossible or expensive to collect the necessary information for decision making. For example, investment projects risk assessments are usually used to forecast sales, costs, prices, etc. However, for an adequate risk assessment sufficient information is required in order to formulate hypotheses about the true probability distributions of key parameters of the project. In such cases, the missing data are replaced with values obtained in the process of imitation experiment (i.e. generated from the computer). In solving many of the tasks of financial analysis are using models containing random variables whose behavior is uncontrollable by the decision makers. Those models are called stochastic. Use of simulation allows us to draw conclusions about the possible outcomes based on probability distributions of random variables. Stochastic simulation method is often called Monte - Carlo.

5. Simulation method Monte Carlo

Simulation model called "Monte Carlo" is based on efficiency calculation in all possible combinations of variable interact. The calculations carried out with the PC, simulated spins of the gambling roulette. After a sufficient number of imitations while modeling different cash flows, project suitability is determined by certain criteria. Simulation is a tool of financial planning that shapes certain cases. When simulation is used for risk assessment, it requires an assessment of the likely distribution of assessed financial indicators (NPV, IRR, IP, PP, etc.). The computer program is given a large amount of random input variables for each interval and any iteration and it is determined the selected financial indicator for the project.

The model has two important advantages [1], [2], [4]:

– It gives a description of the various options and in some cases an alternative to modify the decision depending on the specific objectives.

– Overcomes the natural technical difficulties, large amount of needed information is readily available or gathered and the validity of used principles and relationships has been demonstrated in other proposals too.

In the present work, based on data from the bellow example a simulation analysis of the risks of an investment project is made.

Example 1

The company is considering an investment project for production of "A". The preliminary analysis of the experts identified three main financial parameters of the project and defined the limits of their modifications (Table 1). It is accepted that other parameters are constant values (Table 2).

Table 1.

Variable parameters of the project to manufacture a product "A"			
	Scenario		
Indicators	**Pessimistic**	**Optimistic**	**Possible**
Unitary Capacity production - Q	150	300	200
Price per specimen - P	40	55	50
Variable costs - V	35	25	30

Table 2.

Constant parameters of the project to manufacture a product "A"	
Indicators	**Most - possible values**
Fixed costs - F	500
Amortisation - A	100
Income tax - T	60%
Discount rate - r	10%
Project Duration - n	5
Initial investment - I_0	2000

The first stage of analysis by Monte Carlo algorithm is to determine the dependence of the resulting indicator depending on the input from the original. The output indicators are usually represented by one of the efficiency criteria - NPV, IRR, PI [7], [10].

Suppose that the criterion used net present value of the project

$$NPV = \sum_{t=1}^{n} NCF_t /(1 + r)^t - I_0 \qquad (1)$$

NCF (Net Cash Flow) - net cash flow in period t.

By subject, the value of the discount rate r and the initial investment I0 are known and are considered constants during the project (Table 2). In order to simplifies the calculations assume that the flow generated by the project NCF payments for each period is equal and is determined by the following ratio:

$$NCF = [Q*(P - V) - F - A]*(1 - T) + A \qquad (2)$$

The next step in the assay consists is in choosing the law of probability distribution of the key variables.

In the example key variable parameters are variable costs - V, Unitary Capacity production - Q and Price per specimen - P. The price ranges of possible changes to various parameters are given in Table 1. The remaining variables - the discount rate R and the initial investment I0, are constant - Table 2. In the calculation it is assumed that all the key variables have normal probability distribution.

The third phase - the simulation can only be done by a computer equipped with special software. Therefore we get to know by means of MS EXCEL to automate its implementation.

6. Simulation modeling using the functions of MS EXCEL

Conducting experimental simulation in MS EXCEL can be achieved in two ways - using built-in functions and with using tool random number generator (RNG), added in 'Data Analysis' (Analysis Tool Pack). In the present paper main attention is focused on the technology to conduct simulation and subsequent analysis of the results using the tool RNG. Using built-in functions make sense only if the probabilities of all values of the random variable are the same. To simulate the values of required variable, you can use mathematical functions RAND () or RANDBETWEEN. The format of these functions is shown in Table 3.

Table 3.

RAND( )	Returns an evenly distributed random number greater than or equal to 0 and less than 1. A new random number is returned every time the worksheet is calculated.
RANDBETWEEN(bottom,top)	Returns a random number between the numbers you specify. A new random number is returned every time the worksheet is calculated. If this function is not available, and returns the #NAME? error, install and load the Analysis ToolPak add-in.

If you set analogous formulas for the variables P and V, and a formula for calculating NPV and copy them the required number of times you can get distribution probability, using a variety values of input. Then, using statistical functions can easily to calculate the parameters of the distribution and to make probabilistic analysis [4], [5], [6], [9].

We will use the proposed approach for the solution of Example 1.
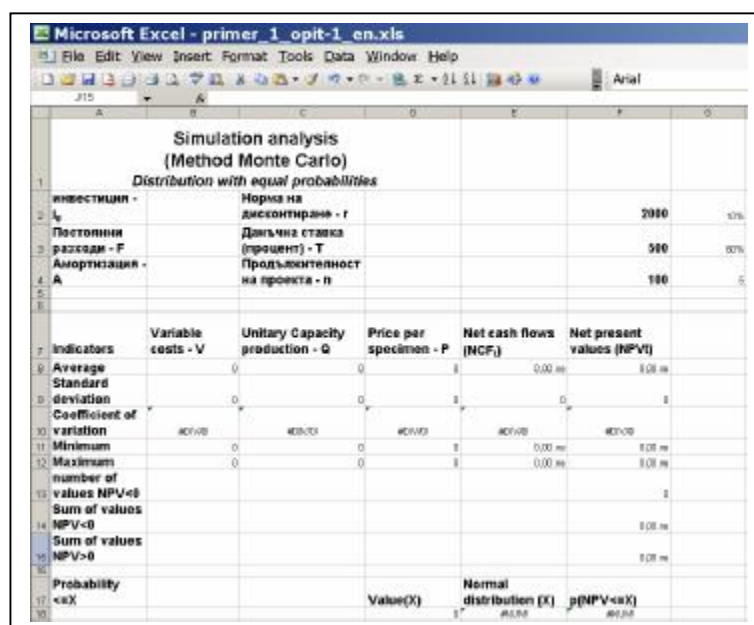
Fig. 1. The experiments performed the results by analysis

Fig. 2 shows a diagram of the distribution of values of key parameters V, P, and Q, built on the basis of 100 simulations. It is easy to

see that the variance of the values of the three casual nature parameters confirms adopted earlier hypothesis of independence.
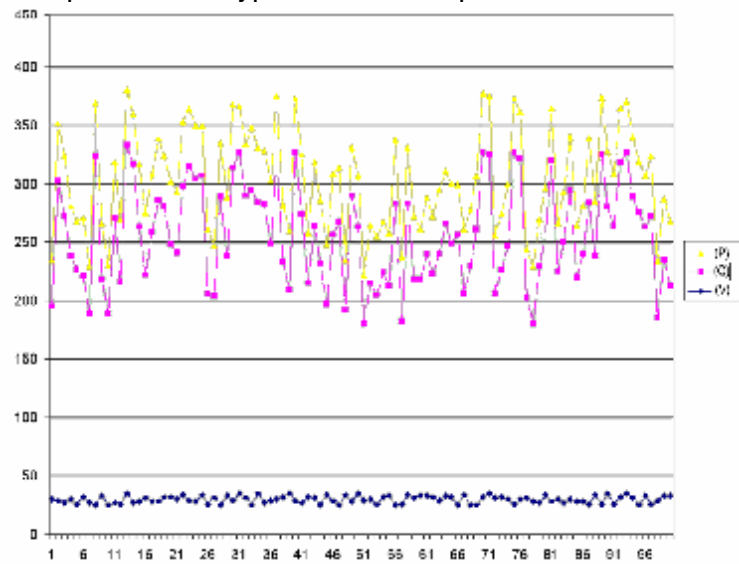
Fig. 2. Distribution of values of the parameters V, P, Q

For comparison, the Fig. 3 shows the distribution of values of key parameters NCF and NPV, built on the base of same 100 simulations.
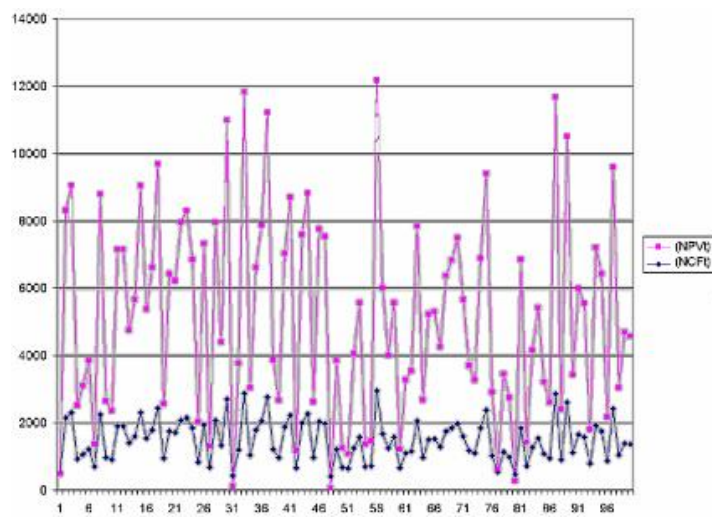
Fig 3. Distribution of parameter values NCF, NPV

Conclusion

Monte Carlo methods are especially useful for simulating systems with many coupled degrees of freedom. They are used to modeling phenomena with significant uncertainty in inputs, such as the calculation of risk in business. Monte Carlo methods are used in finance and mathematical finance, to calculate and analyze instruments, portfolios and investments by simulating the various sources of uncertainty affecting their value, and then determining their average value over the range of resultant outcomes. This is usually done by help of stochastic asset models. The use of the 'Monte Carlo Simulation' allows evaluating investment projects in their dynamic aspect. In many cases, future decisions are influenced by the actions performed in the present. Too often decisions are made without to be taken into account their long-term effects. As a result, decisions that initially appear logical may make decision makers in disadvantageous position in terms of future opportunities. Use of MS Excel for these calculations, makes it easy and accessible for risk assessment and decision support in the investment process. The understanding the risk and uncertainty means understanding most of the problems of investment management. Against this background the simulation is an expression of a specific progress in assessing the probabilities when various factors changed.

Reference list

1. **Славов, З., Брусева, М.** Анализ на риска при вземане на финансови решения. ВСУ, Варна, 2011
2. **Брусева, М.** Управление на риска. ВСУ, Варна, 2010
3. **Славов, З., Д. Дочев.** Методи за измерване на инвестиционния риск. Научен алманах на ВСУ – книжка 7, ISSN 13119222, 2009.
4. **Асват Дамодаран.** Инвестиционная оценка. Инструменты и методы оценки любых активов. Изд. 'Альпина Бизнес Букс', 2004, ISBN 978-5-9614-0802-7
5. **Брусева, М.** Методика и софтуерен продукт за разработване и оценка на инвестиционни проекти за малки и средни предприятия. Варна, 2003
6. **Смирнова, Е. Ю.** Техника финансовых вычислений на Excel. Россия, Санкт-Петербург, 2003 – http://www.cfin.ru/finanalysis/smirnova/
7. **Лукасевич, И. Я.** Анализ финансовых операций. Методы, модели, техника вычислений. Изд. 'Финансы', Москва, 2001.
8. http://pages.stern.nyu.edu/~adamodar/
9. http://www.risk24.ru/
10. http://www.ifri.ch/links_ifri.htm

Assoc. Prof. eng. math. Mariya Nikolova Bruseva, PhD
Varna Free University, Chajka resort, Varna 9007
Tel: +359 52 359 575, Mob: +359 897 80 55 02
mariya.bruseva@gmail.com

INFORMATION THREATS IN THE INTERNET, SOCIAL NETWORKS AND GAMES

Stanislav Dimitrov, Milena Tulechka, Elena Paunova

Abstract: *Evolution of Internet has created new applications that give us new opportunities for communication. Some of them are the social networks. With the growing of the communication abilities, the threats and the possibilities for attacks are also increasing.*

This paper examines the targets of computer crimes and the different types of malicious software that can infect the users' machines. We also give a description of some methods for reducing the risks associated with these cyber threats.

Key words: *Cyber threats, Social networks, Cloud computing, Mobile devices, online games*

Introduction

In the last few years we are observing a new trend progress in the Internet communication - the social networks. These networks have become a global and cultural phenomenon being an adapted concept of real-life interactions in cyber space. This is possible through Web 2.0 which is a powerful interface that has allowed communication in various social groups. Social networks are constantly progressing new way for doing business giving possibilities for sharing information online and discovering new markets. However these platforms commutation interface is still a weak side. This is assigned and to the security about their normal usage. For destroying the defense of social networks the attackers utilize different tactics and tools for achieving their target [1].

The information threats are a reality even in the online games, which means that the children are not safe while they are playing. The extremely increasing number of the social networks users makes this problem significant. Some aspects of the tablets' and smartphones' information security are included in the present work. In this research some of the information threats are shown and different recommendations for avoiding them are given.

The paper is structured as follows: In the first section there are introduced the most popular and widespread viruses nowadays. The second section shows the most significant threats in the social networks. Then the paper presents the hidden dangers of the online games. In section 4 there are shown some threats for the users of mobile devices and removable storage devices (like USB flash memory sticks). The next

section explains the information threats in the cloud computing. The final section presents a short conclusion.

1. Information threats in the form of viruses

Due to the large usage of the information the losing and stealing of data are also increasing. The data can be lost for any computer devices such as servers, desktop, laptop, mobile phone and etc. Another way for losing information is the human negligence. Nonetheless what the error origin is, if this information falls in hackers hands, they can steal or damage it. That leads to financial losses and reputation spoiling.

Tools for catching and stealing information:

Malware is software for infiltrating or damaging computer system without permission of its owner. These type applications are viruses, worms, spyware, adware and Trojans. They are designed to steal personal information, passwords and distribution spam. For protection from malware Sophos [6] recommends proactive detection technologies that can catch malware code quickly and easy.

According to research of Sophos 14.8 % from malware software is **conficker** worm. That has made it the biggest network threat. A good user defense is the usage of patching security by Microsoft, but it also has hidden frauds that are targeted attacks.

The technology for infecting with **fake antivirus** is: Initially a message that informs user that his computer is infected is appearing. Secondly, the owner gets opportunity to scan his machine with a free antivirus version, but after this lure the user has to buy the fake antivirus program.

The targeted attacks are oriented to large companies. The hackers' motives for these attacks are financial gain and covering new information secretes. This type attacks are done via social engineering as sending e-mail from friend or colleague and the victim opens this e-mail.

Attackers also use flaws of securities and install malware code.

The targeted attacks have done through PDF and image files transferring as HTML code. They can hide malicious code in these files. Hackers use other web-based technologies on Java platform to realize cross-browser, multi-platform technologies [8].

According to official data from SophosLabs [6] an increasing number of targeted attacks is being observed. The phishing method uses spreading of malware software via e-mails.

According to a recent research of Symantec [8] the web-based attacks are also increasing. The stealth attacks are done through TDL rootkits that store their hidden files at the hard disc on partition C and it is starting before the operation system.

16

In order to protect him from these attacks the users have to know the problems of visiting unknown site.

For protection from these malware codes anti-spam software for capture non-target spam can be installed. Another way for defense is limiting services of users.

Drive-by download is a method for attack that uses vulnerabilities in user's browser, browser plug-in, application or operating system.

Ways of contraction victim are:

*) visiting malicious site

*) browsing at legitimate site on that they have a trust and click on dangerous link.

*) hackers had made redirecting to malicious page.

Blackhole is the most popular drive-by malware that give administrator's right of criminals. This kind of malware is used at server-side to generate malicious code.

The product has polymorphous characters and with this is hardly to be detected by antivirus programs.

The infection of the exploited site with Blackhole includes:

*) bot-type malware –Zbot

*) Rookit-droppers are TDL and ZeroAccess

*) Fake antivirus

Protection from Blackhole:

By continuously tracking, monitoring and writing a new address of malware site, the attacker moves sites of new addresses because it is important to have layers of protection.

According to [6] the layers that are tracing for protections are

*) java script –(Mal/ExpJS-N)

*) java exploit components

*) Flash (Troj/SWFExp-AI)

*) PDF (Troj/PDFEX-ET)

2. Social networks

Facebook is the most popular social network that works through profiles and communications. Before the customers can use it they must create personal profile. Each user has a choice what information to provide for his personality. The basic information can include different groups such as: workplace, school or collage, close friends. The communications in Facebook are accomplished through sending personal message, invitation to participate in event, sharing photos, hyperlinks and videos. Facebook is allowing the actors to make their own applications, add modules and play games.

Twitter is an online social network and microblogging service that enable users to send and read text-based posts of up to 140 characters known as "tweets". It has been described as "SMS of the Internet". A special peculiarity of Twitter is that an unregistered user can read the tweets, while registered can post tweets through the website interface, SMS or application for mobile device.

MySpace is the leading social network entertainment destination power by passions of fans such as music, movies. It is the most used social network in the USA.

LinkedIn is a business-related social network site. It is mainly used as a professional network.

Orkut is social network website, which service is designed to help user meet new and old friends and maintain existing relationship. It is the most popular social website in India and Brazil.

Hackers have two reasons for attacks on the Internet: financial gains and a form of protest. The techniques for the so targeted purposes are: social engineering, trust, and fake-accounts. Here is a review on the main attacks and malicious users of social networks.

The **credential robbery** is associated to identification of a fake user in a social network. The target is to defraud users that belong to same social group. For achieving this aim the attackers have to gather information through different tools such as keylogger or spyware. A new way for obtaining data is monitoring of the network traffic between the infected machine and the social network servers. The negligence of users has promoted efficiency of these methods. Another way to accumulating information is appearing when user downloads, clicks some link or executes application.

When creating new social network account, a **fake identity** can be used to obtain trust. After that the fake personality uses this possibility to denigrate the image of a real person. To be a successful fake identity, the attackers have passed thought different steps to accumulate information for obtaining trust. These steps are:
   a) Researching data for the real person at Internet
   b) Direct access to information thought other person
   c) Recovery of document of rubbish
   d) Theft or robbery of documents
   e) Monitoring of postal correspondence
   f) Social engineering techniques

There are two modes **exploitation of trust:** manual and automatic. Through the manual mode attackers communicate with victim and make behave of trust. Before this the user is being well observed. The

18

automatic mode uses different instruments to collect trusted identity. After that it installs some type of malware software.

Bulling is an aggressive behavior manufactured by the usage of force to affect other customers. This method has been used before ages. Nowadays there is influence on victim via new technologies and this action is known as **cyberbulling**. Their range of react is larger than the range of the bulling. Through cyber harassment the aggressor doesn't see victim and hence have smaller influence on the people.

Phishing is a way of attempting to acquire information directly to attacker. These data is collected by sending spoof e-mail or instant messaging. To be successful the phishing are combined with different techniques of identity robbery, fake identity and other. Also programs for automatically sending messages and e-mails are used [3].

3. Threats in the computer games

According to pan-European study over 80% of students use the Internet for activities related to school [2].

The most serious and common threat of the games is the theft of information that can be used for illegal access to the bank account of the victim. Accounts in some gaming platforms contain information about the user's bank account. This information is the main subject to attack by cyber criminals in the recent years [9].

3.1. Paid free games and inappropriate content

The majority of games in social networks (and not only there) are advertised as completely free, but they contain parts that are available after paying with real money. The payment is usually done by bank transfer, paid SMS or online payment systems (e.g. PayPal). Examples of these games are the popular FarmVille [15] and other games of the Zinga Company in the social network Facebook [16], the games of the Upjers (Molehill Empire [13], My Free Farm [14], etc.) distributed on the Internet. This is not a direct threat, but it misleads the consumers [7].

The inappropriate content in the plot of the game must also be taken into account. In many games major roles play violence, sexuality, gambling and the distorted reality. These aspects are undesirable, especially for children.

3.2. Software robots

The software robots (so-called "bots") are programs or methods that simulate user actions in a game or other application. Their purpose is to acquire an unfair advantage, but not an identity theft or attacking bank accounts. These programs can constantly collect resources in a game, even during the time when the user is away from his computer. In the MMORPG (like World of Warcraft, [17]) "bots" can affect the play of other

players and the price of game currency purchased with real money [7]. Another way that bots give unfair advantage of consumers who use them is that their actions are much more accurate than the reactions of real people, which is particularly useful for direct competition between players such as the fight between avatars.

Several methods for detecting bots have been developed. They use tracking the way, the sequence of movements, intervals between pressing key and generated Internet traffic [9].

3.3. Other forms of fraud in the games

Fraud in online gaming is considered as any action that's used to gain advantage or achieve a goal if, according to the rules or the discretion of the operator, the advantage is unfair to the other players or the goal should not have been achieved.

Pritchard [4] divides the fraud into 6 categories as follows:

– Reflected acquisition: using a computer program to replace the human reaction to obtain better results in action games.

– Dictatorial customers: using dishonest clients to send modified commands to honest customers who execute them blindly.

– Information disclosure: using access or visibility of hidden information from risk client software.

– Compromised servers: Changing the server configuration to earn an unfair advantage.

– Bugs or flaws in the design: Using bugs or errors in the game software.

– Weakness of the environment: using specific hardware or operating conditions.

3.4. Other threats in the computer games

Many of the online games (the already mentioned FarmVille, Molehill Empire, etc.) are played by millions of mouse clicks. Clicks to interact with each object in the game are simply mechanical but they effectively hold the attention of the person in one place [7]. The players lose time, which can be spending in more useful activities. In recent years information came about some cases where consumers have lost their jobs or have been abandoned by an intimate partner because of their addiction to one or another game. Perhaps the most striking example is a gamer that died as a result of several days constantly playing a game [12].

For a quick and easy progress in the games on social networks, friends playing the same game are needed. The aim is to exchange gifts or other assistance, while even more people are fascinated to play this game [7].

Long time in a sitting position can cause incorrect motor growth in children and obesity in all users.

Although modern technology offers screens with lower levels of electromagnetic radiation, the continued sitting in front of them means a greater amount absorbed radiation. This can lead to vision problems, and many studies have shown that radiation causes more serious health problems.

The games are also used by people seeking a real appointment with users and having criminal intent. In this case, the most vulnerable are children who may become victims of pedophiles or different sects.

3.5. Recommendations for avoiding threats

To ensure the safety of their children, parents should educate and inform them for the dangers prowling them in the virtual world, especially from the strangers. It is also efficient to use parental control software [5].

4. Mobile devices and removable media

Removable media are devices that are very convenient for storing information; however there are often threats in the stored data. Due to the small size they can be easily stolen or lost.

According to survey of Sophos Australia 66% of 50 USB devices lost in public were infected with malware. At device it was able to be found information about the former owner, his/her colleague or friend.

Protection when using removable media

User should encrypt all personal and business data before storing on removable media so it can't be accessed in case those devices are lost. Other protection from spreading malware code via removable device is disabling Windows autorun when insert CD or USB device and it will not start automatically.

Mobile devices

According to the marketing analyst International Data Corporation the market of tablets and smartphones has increased continuously at the last years. They analyzed that the use of mobile device had rose with 300% in the second quarter of 2011 compared to the same period in 2010. This trend has changed technologies and methods for defense of the business regarding to information security. With the increasing use of mobile device for working with theirs application or checking received messages, accomplishment of banking and other transactions the attacks on them are growing.

The mobile malwares are fake online banking applications to steal customer credentials, intercept banking authentication token code via SMS and robbing bank accounts. The criminals are infecting victims by publishing malicious changed applications at the Android Market, e.g.

popular games such as "Need for speed" and "Angry Birds". They were downloaded over 10,000 times only for a day. Other problem for personal computers is careless using from the user's side. Strategies for protection of malicious hackers' attacks are strong password, data encryption, regularly updating patch of software and educating the users.

It is recommended to customers to use security software for their PCs, smartphones or a tablet, which is updated regularly, and not to store personal information on the gaming device or to open random links. To avoid malicious applications, the users should only install such ones from the official stores for their platforms. It is also good to check their sources and consumer reviews. You can do a quick search on the name of the manufacturer to see if there was anything suspicious about the application.

5. Cloud computing

The cloud computing is a new technology for data storage, access and communication to this data over the Internet. The usage of cloud computing is orientated towards creating scalability, flexibility and lower cost. An advantage of the technology is that the services can combine public, private and hybrid environment for data storage.

Threats at cloud computing are similar and include every threat mentioned above. The easiness of registering is precondition for attacks by cyber criminal. There is a variety of ways in which in-the-cloud capabilities can be misused such as launching dynamic attacks points, password and key cracking.

Interfaces that are used by customers must have extremely secure, authentication, access control, encryption and activity monitoring mechanisms. The malicious insiders also are an important problem for the security of cloud computing due to the fact that providers don't know who are the people using their services [11].

The losing or stealing of data can happen when the deleted files are without back-up, not encrypting data and there is an unauthorized access to them. The essential threats are connected to the security because the providers are focusing to functionalities and they don't use appropriate software update, intrusion prevention and firewall. The users also provoke problems such as click links in e-mail messages, visited fake web sites, download malware. The hackers' attacks are against software operation systems. The last trend of their goals is pointed to web browser vulnerabilities [10].

Protections from loss of data at cloud computing

The first protection is encrypting data before storing it in the cloud. The second protection is carefully reviewing and choosing the cloud provider about security measures, back-ups and failover.

Conclusion

Internet has already entered all spheres of our everyday life - work, education, entertainment, communication. This fact, together with the development of new and more convenient devices increases the number of the online users. In parallel, however, the number of harmful people, who take advantage of the new technologies and the reduced vigilance of Internet users, is increasing. Numerous malicious applications are developed, and some of them are built in ordinary programs and files. It is important for the users' protection to develop own habits and to support their children in the protection against such information threats.

References
1. Anchises M.G. Security aspects and future trends of social networks www.icofcs.org/2009/ICoFCS2009-PP09.pdf
2. ARC FUND Presentaion – Център за безопасен интернет, http://safenet.bg/fileSrc.php?id=20569
3. Learte Peotta de Melo, Edna Dias Canedo, Robson de Oliveira Social Network: Security and Privacy http://www.scribd.com/peotta730/d/49166691-2010-Social-Networks-Security-and-Privacy-Peotta-Et-Al
4. Pritchard, M. *How to Hurt the Hackers: The Scoop on Internet Cheating and How You Can Combat It*, Information Security Bulletin, February 2001.
5. Sentonas M. *Cyber threats a reality in online gaming,* ABC Technology and Games, 20 Dec 2011, http://www.abc.net.au/technology/articles/2011/12/20/3394645.htm
6. Sophos Security Threat Report 2012 http://www.sophos.com/medialibrary/PDFs/other/SophosSecurityThreatReport2012.pdf
7. Steinberg, S. *Social Games: 5 Growing Threats to Watch*, Posted November 15, 2011, http://www.industrygamers.com/news/social-games-5-growing-threats-to-watch/
8. Symantec Internet Security Threat Report http://msisac.cisecurity.org/resources/reports/documents/SymantecInternetSecurityThreatReport2010.pdf
9. SYSSEC Preliminary Report on Social Networks Security, *A European Network of Excellence in Managing Threats and Vulnerabilities in the Future Internet: Europe for the World* http://www.syssec-project.eu/media/page-media/3/syssec-d5.2-SoA-SocialNetworkSecurity.pdf
10. Top 10 cloud computing threats http://www.altiusit.com/files/blog/Top10CloudComputingThreats.htm
11. Top 7 threat of cloud computing http://www.net-security.org/secworld.php?id=8943

12. http://www.24chasa.bg/Article.asp?ArticleId=792653 *Геймър умря пред компютъра, играл три дни без почивка*, Публикувана: сряда, 23 февруари 2011
13. http://www.molehillempire.com/
14. http://www.myfreefarm.co.uk/
15. https://www.facebook.com/FarmVille
16. https://www.facebook.com
17. http://eu.battle.net/wow/en/

About the authors
M. Sc. Eng. **Stanislav Dimitrov Dimitrov**, PhD student, Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, 02/979 2774, sdimitrov@hsi.iccs.bas.bg
M. Sc. Eng. **Milena Kirilova Tulechka**, Military Medical Academy, m_tulechka@abv.bg
M. Sc. Eng. **Elena Nikolaeva Paunova**, Res. Assistant, Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, 02/979 2774, eli_np@abv.bg

EFFICIENCY CRITERIA FOR USE OF SATELLITE SYSTEMS IN MOTOR TRANSPORT

I. Gonzalez – Ortiz, V. Boicov

General

Today in the conditions of the increasing competition among the motor transportation companies it is necessary to provide quality of service demanded by consumers and efficiency of transport operations. This task can't be solved without introduction of information technologies in management of productions, i.e. application of information systems and program complexes of the analysis, planning and support of making commercial decisions.

In recent years in the market of the monitoring system and monitoring of transport use there is the increasing interest which help to supervise remotely vehicles and everything that occurs on their route. The most important fact is that as a result the owner of vehicle fleet receives an essential economy of finances. On the average run the vehicle fleet decreases for 20-40 %, fuel consumption - for 20-60 %, and general charges of vehicle fleet - for 20-30 %. The investments, as a rule pay off within the first year, and sometimes already in the next months.[2]

Thanks to introduction of information systems and the technologies in the motor transportation companies there is the possibility of automation of standard operations in transport processes. As a result the logistics has become a dominating form of the organization of movement of goods in the market of transport services. From the point of view of logistics the requirements to information technologies should be the following:

§ to analyze a situation and behavior of cooperating elements of system in real time,

§ to provide monitoring and diagnostics of administrative processes in a dynamic mode,

§ to model real actions and events,

§ to predict and forsee critical situations.

Transport – logistic activity deals with the solution of a number of tasks. In what terms cargo should be delivered? What expenses will be necessary for cargo movement? What route should be chosen? How to provide safety of cargo? And so on. If these sorts of tasks are solved constantly, final losses at the expense of inefficient decision-making can

essentially affect the general expenses (and, respectively, the firm profit). [3]

1. Assessment of efficiency of transport process

Key question of a choice of the most rational variant of transportations is the assessment of efficiency of a transport process. The choice of criterion of efficiency depends on specific conditions of transportations and a solution of a task. We can distinguish the local (private) and generalized (complex) criteria of efficiency.

Local criteria of efficiency apply, if compared to the options of transportations differ from one separately taken indicator. So, introduction of hour schedules of transportations excludes idle times of cars in a line. In this case efficiency of compared options of transportations can be estimated by one indicator: duration of idle times of the car in loading and unloading points. Also the use of cost assessment of idle times of transport is possible. Introduction of rational routes of transportations provides reduction of a single run. In this case distinction in compared options can be estimated by reduction of empty run of cars or other indicators connected with empty run:

- run efficiency;
- general run;
- fuel consumption etc.

Complex indicators of efficiency apply when certain actions at the same time change some characteristics of a transport process. For example, replacement of a rolling stock leads to change of such parameters, as:

- loading capacity;
- idle time of loading and unloading;
- specific fuel consumption;
- depreciation charges, etc.

In this case any private criterion is not enough.

As local indicators of efficiency we can use the following technological parameters of transport process:

- average distance of transportation;
- zero run;
- empty run;
- total loading capacity of cars;
- average efficiency of loading capacity;
- total idle time of cars;
- need for cars;
- the ton hours spent for performance of the set volume of transportations;

26

• general time for performance of transportations;
• timeliness of delivery;
• cargo cost on a route;
• speed of delivery of cargo;
• size of losses of cargo on a route;
• safety of cargo.

Number of local indicators of efficiency of transport process also includes the following :
• power consumption;
• material capacity;
• material capacity of transportations. [2]

2. Transport expenses

Transport expenses - a part of transport and procuring expenses; costs of production transportation from places of production to the direct consumers, carried out by both transport of the general use, and own transport. They include payment of tariffs of transport and other collecting transport organizations, charges of own transport, cost of cargo handling works, escort of cargoes, a difference between the actual expenses on payment of tariffs of transportation and the sum compensated by suppliers to the supplying and marketing organizations (taking in regard the average distances of transportations)

Level of a gain of expenses depends on the following variables.

1. The relation of the size of manufacturing enterprise to the capacity of the served market. If the manufacturing enterprise satisfies only a small share of a market demand, it can increase considerably sales volume without expansion of geography of sale.

2. Pricing method. The transport expenses considered by the producer, increase with the growth of volume of production when the price level in all markets will be identical or when the price in more remote markets will be established by the competitors possessing advantages of a site of their production in relation to sales markets. Such conditions are available for many companies.

3. Geographical structure of transport expenses. Usually freight tariffs grow more slowly, than remoteness. The growth rate of the expenses connected with the delivery of production on additional 100 km less, the less transport expenses are connected with the amount of production.

4. Geography of placement of customers. If buyers are evenly distributed on the region territory, at increase in number of clients transport expenses will increase to a lesser extent, than other things being equal as costs of delivery depend on delivery radius while sales

volume - from a radius square. If the density of distribution of customers sharply decreases in process of removal from the market, transport expenses at increase in volume of deliveries can increase in much bigger degree.

5. Relation of production expenses to physical volume of a unit of production. In transit of the volume goods having the low price, such as sand or beer bottles, transport expenses dynamically grow with increase in distance of delivery. For the compact and expensive goods, type of integrated schemes and tools, transport expenses grow slowly. [2]

3. New standards of service

Regularly carried out researches of preferences of economically active groups of consumers show gradual change of criteria according to the service choice. Undoubtedly, the low price and high quality of the goods are still the most important in this list of criteria, however they are not so simply desirable characteristics, but more likely an obligatory standard of competitive service. Today consumers even more often pay attention to additional criteria, such as delivery periods, possibility of receiving the ordered goods during accurately stipulated time, and also high-quality information support of the process of implementation

For the head of the company which vehicle fleet totals in ten, and even more than tens vehicles, it is difficult, and in most cases - it is impossible to check a real location of all vehicles, existence of malfunctions, fuel consumption, no-purpose vehicle fleet uses («the left flights»), observance of routes, untimely delivery of cargoes etc. Owing to poor control and abuses of the company the additional financial losses, the spoiled image and possibility of loss of clients can threaten.

The system of satellite monitoring of transport will help to solve all aforementioned problems.

4. Satellite monitoring of transport

Satellite monitoring of transport is the system of satellite monitoring and management of the mobile objects, constructed at the basis of use of modern systems of satellite navigation (GPS/GLONASS), the equipment and technologies of communication (GSM/GPRS), computer facilities and digital cards.

Principles of work of satellite system of monitoring and transport scheduling are the following; onboard equipment information is installed on the car with the help of which the information is transferred through the GSM/GPRS networks and the Internet to the computer of the dispatcher, it is also displayed on electronic district maps, and reports in the form of schedules and in tables are formed. Location, passable

routes, speed and movement direction, time and place of stops, fuel consumption, operation of sensors, crossing of inspection zones are specified.

For more information on a vehicle the additional sensors connected to GPS or GLONASS to the controler, are established, for example:

- fuel consumption sensor;
- sensor of load of the HARDWARE axis;
- the fuel level sensor in a tank;
- the temperature sensor in the refrigerator;
- the sensors fixing the fact of work or idle time of special mechanisms (turn of an arrow of the crane, work of a concrete mixer), the fact of opening of a door or a cowl, the fact of presence of the passenger (taxi);
- management of ignition;
- monitoring of a charge of the battery of the accumulator;
- disturbing button;
- hitch control;
- control of load of an axis.

The GPS system of monitoring of transport will provide new possibilities:

1. To reduce operational costs of vehicle fleet at the expense of increase of efficiency of control of moving of motor transport, optimization of routes.

2. To reduce fuel consumption at the expense of effective control of its use.

3. To increase labour productivity, to improve discipline of the personnel, to develop effective system of payment and motivation of work of drivers at the expense of effective control of performance of tasks, routes and schedules.

4. To increase management efficiency and to improve quality of the account as a result of receiving reliable and operational information about motor transport and cargo site, run, stops, time on routes, speed of movement, fuel consumption, performance of routes, driving schedules etc.

5. To increase safety of transportation and to provide safety of cargoes thanks to control of access to motor transport and cargo, monitoring of speed of movement and a deviation of vehicles from the set routes.

6. To reveal no-purpose use of motor transport operatively.

7. To reduce probability of the risks connected with excess of speed that allows to minimize influence of a human factor and to promote increase of labour productivity of drivers.

8. To increase efficiency of transport logistics at the enterprise. The system gives opportunity to develop optimum routes and traveling tasks and to carry out their operative adjustments.

9. To cut down expenses on mobile communication at the expense of economy on telephone conversations of the dispatcher with the driver.

10. To integrate the obtained data into a uniform enterprise management system.

11. To increase quality of servicing, to reduce cost of services, to increase competitiveness, growth of profit.

12. To receive the detailed reports about non-standard and the emergencies demanding operative reaction. [1]

5. GPS — global system of determination of coordinates

GPS (Global Positioning System — global system of determination of coordinates) — the satellite search engine made of set of 24 satellites, placed into an orbit the American Ministry of Defence and the land stations of tracking united in the general network. The global system of determination of coordinates (GPS) was originally intended for the military purposes, but in the eighties the XX century, the government made system available to civil use. The global system of determination of coordinates works in any weather conditions, worldwide, 24 hours in a day. Any restrictions on use of system of determination of coordinates doesn't.

Satellites of system move on an exact orbit with a cycle time of 12 hours and hand over information on the earth. GPS receivers accept this information and, using a triangulation (breakdown on triangles), calculate exact location of the user. In essence, the GPS receiver compares time transferred by the satellite over time when this time was sent. The time difference speaks to the receiver about how far there is a satellite. Having measured such distance to several satellites, the receiver can define position of the user and show it on the electronic card of the module (block).

The receiver (GPS navigator) should be adhered to signals, at least, three satellites for determination of two coordinates (width and a longitude). Having four or more satellites under review, the receiver can define three coordinates of the user (width, a longitude and height). As soon as position of the user will be defined, the system can calculate other information, such as speed, the course, the passable distance, distance to a point of appointment, sunrise and decline time etc.

6. Effect from GPS/GLONASS system use

Run reduction	20 – 25%
Economy of fuels and lubricants	20 – 25%
Reduction of an idle time of transport	10%
Increase in an turn - over of transport	15%
Decrease in expenses on transport operation	20 – 25%

The system pays off from 3 to 18 months.

Scientifically – technical progress causes transition from out-of-date, conservative methods of regulation and optimization of transport transportations to more modern and progressive, providing reductions of time of deliveries and minimization of expenses. [1]

Conclusion

The actuality of transport problems proves to be true by the fact that about 50 % of all expenses for logistics are connected with transport expenses. No economy can develop progressively, if the effective functional model of transport system isn't created. Transport in many cases acts as the catalyst of increase of level of activity of economy. Connecting production and consumers, it allows to expand production and consumption scales.

Carried-out polls show that use of modern systems of positioning for transport, and also monitoring systems of motor transport are economically justified. Therefore it is obvious that economic effect of their application in the large companies which are carrying out thousands deliveries in a day, repeatedly exceeds costs for its acquisition, introduction and service support.

As foreign experience testifies, high-quality "jump" in the transport sphere can be reached only at the expense of use of new technologies that ensure processes of the transportations meeting modern requirements and high international standards.

References
1. Десятник, В. В., Лысенко, Д. Э. Методы мониторинга транспорта в логистических цепях. Бизнесинформ №6, 2010
2. Дорофеев, А. Тренды развития транспортной логистики. Логистика.№2, 2011

3. Воропаев, А. Комплексное решение задач транспортной логистики. АвтоТрансИнфо №1, 2011

Authors:
Irina Gonzalez – Ortiz, Docent, Mg. Sc.
V. Boicov, Associate Professor, Dr.Sc.

Information Systems Management Institute
Department Of Natural Sciences And Computer Technologies
Ludzas 91, LV-1003, Riga, Latvia,
e-mail: irinmart@inbox.lv, phone:+371 29836562
e-mail:boicov@yahoo.com, phone:+371 29857487

INFORMATICS IN NANOTECHNOLOGY RISK MANAGEMENT

Nadežda Kņiga

Abstract: Not developed criterias of risks estimation for the nanogoods manufactures lead to a decrease in the profit of the enterprise and represent potential risks for health, environment and social safety. Therefore this article is intended: to provide information of negative aspects of nanomaterials science and public perceptions of nanotechnology risks; to do review of existing risk management principles and new risk management approaches for nanotechnology; to suggest Information Technology Solution for optimal business decisions making.

Key words: nanomaterials, nanotechnologies, nanothinking, nanoeducation, nanorisks, regulation, marketing research.

1. Potential dangers of nanotechnology

1.1. The impact of nanomaterials in industry and society

What is nanomaterials science and Nanotechnology? Nanoscience and nanotechnology are based on the control of the structure and function of materials on the nanometre scale, i.e. on the scale of one billionth of a metre. The gateway to nanoscience and nanotechnology has been opened more than 100 years ago, when W.C. Röntgen discovered the x-rays which allowed us to unravel the nanoscale structure of matter and when Planck, Heisenberg, Schrödinger, and Einstein developed the language of the nanocosmos: quantum mechanics. [2]

Only few industries will escape the influence of nanotechnology. Faster computers, advanced pharmaceuticals, controlled drug delivery, biocompatible materials, nerve and tissue repair, surface coatings, better skin care and protection, catalysts, sensors, telecommunications, magnetic materials and devices – to name but a few of the areas where nanotechnology will have a major impact. In effect, nanotechnology is a radically new approach to manufacturing. It will affect so many sectors that failure to respond to the challenge will threaten the future competitiveness of a large part of the economy. [2]

Advanced materials are a prerequisite for all major research and development areas and for all key technologies ranging from information and communication, health and medicine, energy and environment, to transport and space exploration. [2]

Since applications of nanotechnology will quickly penetrate all sectors of life and affect our social, economical, ethical and ecological

activities, the general public's acceptance is compulsory for further developments of in the field of nanotechnology and its applications. This acceptance will be influenced by the low level of public awareness of many innovations in science, and especially, in nanotechnology. This is mainly due to the unpredictability of their properties at the nanoscale and the fragile public confidence in technological innovation and regulatory systems. [3]

1.2. Risks of nanotechnology

On the analysis basis of the barriers the following risks are formulated:

§ Technological,

§ Economic,

§ Social,

§ And other risks of market operators.

The basis risks are – risks of explosion, risks of a person health and commercial risks which can arise now.

The secondary – the risks, which display occurs indirectly and it is prolonged in time: ecological risks, political risk and risks of safety, risks of rupture in formation, transboundary risks, and risks of public perception, public structure and development of the person.

Potential risks for health:

• *Toxicology of engineered nanomaterials and by-product nano-particles:*

Ability to enter into the human body

Impact of biochemical processes

• *Harmful effects of "beneficial" nanomaterials ingested in the human body:*

Impact of nanoparticles used to destroy cancer cells on the whole of the human body

Photochemical reactions of nanoparticles used in sunscreens and cosmetics

Potential risks for the environment:

• *Toxicology of engineered nanomaterials and by-product nanoparticles:*

Accumulation and transportation in water, soil, and the atmosphere

• *Adverse impact of "beneficial" nanomaterials for the "food chain":*

Further impact of nanoparticles transported or transformed via micro-organisms such as bacteria and protozoa

Potential risks for safety, security and ethics

• Invasion of privacy

• Spread of spying sensors

34

• Nanorobotics or other bio-nanotechnology ambitious applications. [2]

Risk management of nanotechnology is further challenged by the broad range of technology and products encompassed within the term "nanotechnology". Risk management of nanotechnology must take into account public perception about the potential risks and benefits of nanotechnology and the growing public demands for regulatory oversight.

Currently, the lack of reliable and accurate technical data on the topic provides fertile ground for both nanotechnology proponents and sceptics to make contradictory and sweeping conclusions about the safety of engineered particles for human health and ecosystems. It is therefore imperative to carry out investigations devoted to the development of a comprehensive understanding of the properties, interaction, and fate of natural and anthropogenic nanoscale and nanoengineered materials in human health and environment. [2]

2. Existing risks management philosophy

Risk management of nanotechnology is challenged by the enormous uncertainties about the risks, benefits, properties, and future direction of nanotechnology applications. Because of these uncertainties, traditional risk management principles such as acceptable risk, cost–benefit analysis, and feasibility are unworkable, as is the newest risk management principle, the precautionary principle. [3]

This paragraph analyzes the applicability of traditional risk management principles and new approaches based on the precautionary principle to nanotechnology, and finds these available approaches to be inadequate and unworkable. [3]

2.1. Traditional Risk Management Principles
The three most common traditional models for risk management of hazardous agents are:
§ Acceptable risk,
§ Cost-benefit analysis,
§ Feasibility (or best available technology).
Acceptable risk approaches rely on risk assessment to describe the risks of an agent, and then seek to reduce risks to levels that are socially acceptable. Current understanding of nanotechnology risks is too uncertain to permit meaningful risk assessment, and is likely to remain so for some time. [3] There are no accepted test methods or validated data that can be used to prepare scientifically credible quantitative estimates of risk of specific nanotechnology applications at this time. [3] Another

complication is the rapid pace of nanotechnology development, which is rapidly outpacing the development of risk assessment for these technologies. [4]

A risk is *acceptable* when:

§ it falls below an arbitrary defined probability

§ it falls below some level that is already tolerated

§ it falls below an arbitrary defined attributable fraction of total disease burden in the community

§ the cost of reducing the risk would exceed the costs saved

§ the cost of reducing the risk would exceed the costs saved when the 'costs of suffering' are also factored in the opportunity costs would be better spent on other, more pressing, public health problems

§ public health professionals say it is acceptable

§ the general public say it is acceptable (or more likely, do not say it is not)

§ politicians say it is acceptable. [6]

A second traditional risk management model is *cost-benefit* analysis or balancing, in which the costs and benefits of proposed risk management options are balanced. [3]

A *cost benefit* analysis finds, quantifies, and adds all the positive factors. These are the benefits. Then it identifies, quantifies, and subtracts all the negatives, the costs. The difference between the two indicates whether the planned action is advisable. The real trick to doing a cost benefit analysis well is making sure you include all the costs and all the benefits and properly quantify them. [7]

Unlike the acceptable risk model, the cost-benefit model has the advantage of considering both the benefits and risks of nanotechnology, which is important given that nanotechnology is likely to present both risks and benefits for public health and the environment. [3] Nonetheless, the cost-benefit model is ill-equipped for managing nanotechnology at this time, given the immense uncertainties about its risks and benefits.

The third and final traditional risk management principle is the *feasibility* or *best available technology* approach. This approach, which requires reduction of risks to the lowest level technologically or economically feasible, has the advantage of not requiring information about risks or benefits. Indeed, the feasibility approach has achieved considerable popularity among policymakers in recent years because it allows circumvention of controversies over risk analysis and jumps straight to reducing risks to the extent possible. [3]

2.2. New Risk Management Principle

In sum, none of the three traditional models of risk management are capable of effectively managing the risks from nanotechnology at this

36

time given the tremendous uncertainties that exist for this emerging technology. This realization has prompted some to advocate application of the most recent entry in the risk management toolbox, the *precautionary* principle. [3]

The *precautionary* principle, which has emerged in recent years as an alternative approach to risk management, is often summarized by the phrase "better safe than sorry."

The precautionary principle recognizes that health and environmental decisions often must be made in the face of pervasive uncertainty, and therefore calls on decision makers to err on the side of safety by delaying new technologies until their safety can be adequately ensured. This requirement is often framed in terms of shifting the burden of proof to the proponent of a technology to demonstrate its safety. Given the massive uncertainty about nanotechnology risks, this technology might appear to be an ideal candidate for application of the precautionary principle. Yet in fact nanotechnology vividly demonstrates the limitations of the principle as a decision-making tool; the precautionary principle too is not a workable risk management model for nanotechnology. [3]

The first problem with the precautionary principle is that it is too poorly defined to serve as a decision-making rule. While lawmakers and proponents frequently cite to "the" precautionary principle, there is no standard text for the principle, and the dozens of formulations that have been suggested differ in important respects. [3]

While the precautionary principle provides a useful general philosophy, and while some application of precaution is certainly appropriate to guide the development of nanotechnology, the precautionary principle itself fails to provide a workable risk management approach. [3]

Yet, simply waiting for these uncertainties to be resolved before undertaking risk management efforts would not be prudent, in part because of the growing public concerns about nanotechnology driven by risk perception heuristics such as affect and availability.

A more reflexive, incremental, and cooperative risk management approach is required, which not only will help manage emerging risks from nanotechnology applications, but will also create a new risk management model for managing future emerging technologies. [3]

3. New risk management model

New risk management model suggest flexible, evolutionary approach to risk "regulation," especially in the immediate, near and medium terms. The approach we suggest is flexible in two senses. Substantively, it draws on multiple approaches to addressing risk, not

only the accepted models of risk regulation discussed above, but also more general approaches to the appropriate handling of risky technologies, such as the notion of product stewardship and the professional ethics of researchers. [3, p 51]

New risk management model is: transparency and dissemination of information; increased level of public trust in those responsible for the development of nanotechnology.

3.1. Regulatory approach

Ayres and Braithwaite employed the construct of a "pyramid" to illustrate their flexible approach to regulation. Most narrowly, this pyramid depicts the spectrum of possible *sanctions* from persuasion and warnings, at the base, up through civil, licensure, and criminal penalties, at the peak available to regulatory agencies. A broader version of the pyramid captures the range of regulatory *strategies* available at the national level from self-regulation at the base, through supervised or enforced self-regulation and other forms of public private interaction in the middle, to standard forms of command-and-control regulation, still with a range of possible penalties as depicted on the original pyramid, at the peak (Figure 1).

Recent scholarship has attempted to apply the regulatory pyramid to nanotechnology. Bowman and Hodge [3] use a pyramid model to argue for a complex regulatory system for nanotechnology (Figure 2). Their pyramid is hexagonal, with six sides that correspond to families of issues including occupational health and safety, environmental protection, product safety, privacy and civil liberties, intellectual property, and international law. In addressing each set of issues, the pyramid suggests that regulators should deploy a range of regulatory options, from "soft law" to "hard law," as called for by Ayres & Braithwaite [3].

Both models, moreover, are designed for advanced nations with highly developed legal systems, in which legislatures and agencies can create, communicate, and utilize a range of regulatory options. Finally, both at least implicitly assume high levels of information and under-standing on the part of regulators. As such, both models may better depict a potential future regulatory end state than an immediate approach capable of addressing the risks of a dynamic technology in an envi-ronment of pervasive uncertainty, as is the case with nanotechnology.
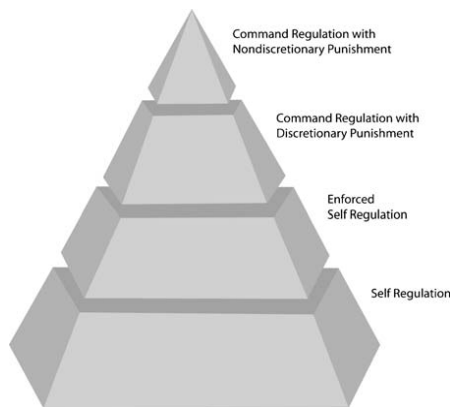
FIGURE 1. AYRES & BRAITHWAITE REGULATORY PYRAMID (ADAPTED WITH PERMISSION FROM AYRES AND BRAITHWAITE [3].
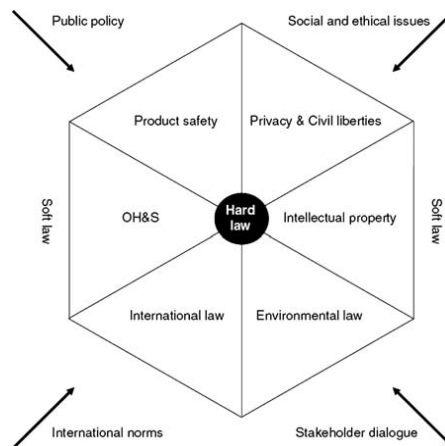
Figure 2. A CONCEPTUAL MODEL FOR THE REGULATORY FRONTIERS OF NANO-TECHNOLOGY. SOURCE: BOWMAN AND HODGE [1].

3.2. Models for marketing research of nanoproducts and nanotechnologies

Since applications of nanotechnology will quickly penetrate all sectors of life and affect our social, economical, ethical and ecological activities, the general public's acceptance is compulsory for further developments of in the field of nanotechnology and its applications. This acceptance will be influenced by the low level of public awareness of many innovations in science, and especially, in nanotechnology. This is mainly due to the unpredictability of their properties at the nanoscale and the fragile public confidence in technological innovation and regulatory systems. Consequently, it is of the utmost importance to educate the public, and to disseminate the results of nanotechnology development in an accurate and open way so that the general public will eventually accept nanotechnology [2].

Marketing research should be based on the objective data of a market, characteristics of groups of the goods and their manufacturers.

The model of marketing research is necessary for its successful carrying out. It details the methods which are required for the information acquisition with which help it is necessary to structure or solve a marketing research problem.

The complexity of the marketing research model in the nanotechnologies is a following: at first it is necessary to develop the uniform criteria for the characterizing manufacture and application of nanoproducts and nanotechnologies. Besides, a rapid development of

39

nanotechnologies advances working out of risk estimation methods, that causes consumers mistrust to nanoproducts and can have catastrophic consequences for their providers and manufacturers.

Risk management of nanotechnology is further challenged by the broad range of technology and products encompassed within the term "nanotechnology". Risk management of nanotechnology must take into account public perception about the potential risks and benefits of nanotechnology and the growing public demands for regulatory oversight.

Currently, the lack of reliable and accurate technical data on the topic provides fertile ground for both nanotechnology proponents and sceptics to make contradictory and sweeping conclusions about the safety of engineered particles for human health and ecosystems. It is therefore imperative to carry out investigations devoted to the development of a comprehensive understanding of the properties, interaction, and fate of natural and anthropogenic nanoscale and nanoengineered materials in human health and environment [2].

3.3. Information Technology Solution for optimal business decisions making

The author of this paper suggests to carry out marketing researches. And problems of these researches are the following:

1) identifying of most commercially effective nanotechnologies and nanogoods and control facilities of consumer behavior;

2) identifying of potentially dangerous nanotechnologies and nanogoods groups;

3) working out of recommendations referring to the regulation of manufacturers responsibility, intermediaries and consumers of nanotechnologies and nanogoods;

4) working out or improving of the risk estimation methodology of nanogoods manufactures.

Research hypothesis: the performance of all research tasks *will allow* to create a public multilevel database including nanotechnologies and nanogoods information which *will help* to make optimal business decisions, *will become* the tool of responsibility regulation of manufacturers and consumers, *will become* the tool of short-term and long-term risks estimation analysis.

This *multilevel database* should:

§ contain the results of nanoproducts marketing research and all actual information what are useful for costumers and manufacturers;

§ have friendly interface and should contain good searching tools.

Object of research is enterprises of the Baltic region which work in sphere of nanotechnologies.

Subject of research is the mechanism of risk management methods at the nanotechnologies enterprises.

Restrictions of research:

§ Territorial. the Baltic region data is used only.

§ Information. Secondary data, existing experts databases, bases of the Latvian Central Statistical Bureau are used for marketing researches (Data mining).

§ Law. Latvia and ES are used. Latvian and European laws are used For the analysis of regulation responsibility tools.

Conclusion

The successful development of such a new risk management approach would not only facilitate the responsible development of nanotechnology, but will create a new precedent that could be used for other emerging technologies of the future. As we look to other technology revolutions looming in the future, including emerging developments in telecommunication technologies, surveillance technologies, genetic enhancement, cognitive sciences, and many others, the need to develop new, better models for risk management becomes all the more urgent. [3]

The author of this paper having made the review of existing traditional and new risk management principles for nanotechnologies, and also having considered offers of new regulation models at the nanotechnology market has come to a conclusion that it is necessary to make the complex marketing analysis of nanotechnologies and nanoproducts of the Baltic States market.

References

1. Breggin LK, Carothers L (2006) *Governing uncertainty: the nanotechnology environmental, health, and safety challenge.* Columbia J Environ Law 31:285–329.
2. Dosch H., Van de Voorde M.H. (2009) *Gennesys White Paper*, Max-Planck-Institut für Metallforschung, Stuttgart, Heisenbergstr. 3, D-70569 Stuttgart,Ch. 10, 310-319.
3. Marchant G.E. & Sylvester D.J. & Abbott K.W. (2008) Nanoethics*. Risk Management Principles for Nanotechnology.*, Springer Science + Business Media B.V.,Ch.2, 43-60.
4. Renn O, Roco MC (2006*) Nanotechnology and the need for risk governance. Journal of Nanoparticle Research* (2006) 8: 153–191, DOI 10.1007/s11051-006-9092-7
5. http://www.portalnano.ru/read/sci/analit/azoev *Г.Л. Азоев: "Маркетинговый анализ рынков нанопродуктов (результаты аналитического проекта)"*

6. 2001 World Health Organization (WHO). Water Quality: Guidelines, Standards and Health. Edited by Lorna Fewtrell and Jamie Bartram. Published by IWA Publishing.
7. http://management.about.com/cs/money/a/CostBenefit.htm   Cost   Benefit Analysis, aka Running The Numbers, By F. John Reh, About.com Guide.

ISMA Graduate School
Высшая школа менеджмента информационных систем
Department of Natural Sciences and Computer Technologies
Ludzas 91, LV-1019, Riga, Latvia,
e-mail: nadezda.kniga@isma.lv,
fax: +371 7241591,
phone: +371 7100690,
mob +371 26426254

ARCH MODELS – USE IN ECONOMICS

Savka Kostadinova and Nikola Naidenov[**]

Abstract:*This paper presents some ARCH models for analysis and prognostication of the economic systems changeability on macro level (inflation and interest), as well as the opportunities for using the spline functions and exponential polynomials for their realization.*
Key words: ARCH, spline, exponential, inflation

After more than 30 years of research on market stability, the central problem of clarifying the *relation between the state of the economy and aggregate financial volatility* still remains unsolved. In his pioneering paper [3], R. Engle proposed a model of the volatility as a combination of *macroeconomic effects* and time series dynamics. Unconditional volatility varies over a long period of time as a result of high volatility of macroeconomic factors such as *GDP, inflation and interest rates.* The low frequency component of volatility is greater when the macroeconomic factors GDP, inflation and short term interest rates are more volatile or when inflation is high and output growth is low. Volatility is higher for emerging markets and for markets with small numbers of listed companies and market capitalization, but also for large economies. Engle's model, known as ARCH model, allows the high frequency financial data to be linked with the low frequency macro data. As a result it will be possible to forecast *the effect of potential macroeconomic events.* This model gives raise to a variety of GARCH models, see e.g. [4] and the references therein. However, the small improvement is obtained due to significant complication of the metod itself as well as the afterward analysis. That is way the ARCH model become a classical tool in the econometrics.

Autoregressive conditional heteroscedasticity (ARCH) model

Conventional wisdom has it that the problem of autocorrelation is a feature of time series data and is a feature of cross-sectional data that play an important role in the analysis of time series ([2]).

Researchers engaged in forecasting financial time series, such as stock prices, inflation rates, foreign exchange rates, etc. have observed that their ability to forecast such variables varies considerably from one time period to another. For some time periods the forecast errors are relatively small, for some time periods they are relatively large, and then

43

they are small again for another time period. This variability could be due to volatility in financial markets, sensitive as they are to rumors, political upheavals, changes in government monetary and fiscal policies, and the like. This would suggest that the variance of forecast errors is not constant but varies from period to period, that is, there is some kind of autocorrelation in the variance of forecast errors.

Since the behavior of forecast errors can be assumed to depend on the behavior of the disturbances u_t, one can make a case for autocorrelation in the variance of u_t. To capture this correlation, Engle developed the ARCH model. The key idea is that the variance of u at time $t(= s_t^2)$ depends on the size of the squared error term at time $(t-1)$, that is, on u_{t-1}^2.

To be more specific, let us revert to the k-variable regression model:

$$Y_t = b_1 + b_2 X_{2t} + ... + b_k X_{kt} + u_t$$

and assume that *conditional* on the information F_{t-1} available at time $(t-1)$, the disturbance term is distributed as

$$u_t / F_{t-1} \sim N[0, (a_0 + a_1 u_{t-1}^2)]$$

that is, u_t is normally distributed with mean zero and variance of $(a_0 + a_1 u_{t-1}^2)$.

Since in previous formula, the variance of u_t depends on the squared disturbance term in the previous time period, it is called an **ARCH(1)** process. But we can generalize it easily. Thus, an **ARCH(p)** process can be written as

$$\text{var}(u_t) / F_{t-1} = s_t^2 = a_0 + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + ... + a_p u_{t-p}^2$$

If there is no autocorrelation in the error variance, we have

$$H_0 : a_1 = a_2 = ... = a_p = 0$$

44

in which case $\mathrm{var}(u_t)/F_{t-1} = a_0$, and we have homoscedastic error variance.

   As Engle has shown, a test of the preceding null hypothesis can be easily made by running the following regression:

$$\hat{u}_t^2 = \hat{a}_0 + \hat{a}_1 \hat{u}_{t-1}^2 + \hat{a}_2 \hat{u}_{t-2}^2 + ... + \hat{a}_p \hat{u}_{t-p}^2 \tag{1}$$

where $\hat{u}_t$, as usual, are the OLS residuals estimated from the original regression model.

   One can test the null hypothesis H_0 by the usual F test, alternatively, by computing nR^2, where R^2 is the coefficient of determination from the auxiliary regression (1). It can be shown that

$$nR^2 \sim c_p^2$$

that is, nR^2 follows the c^2 distribution with p degrees of freedom.

Setting of the problem and the algorithm

   Given a time series $\{y_t\}_{t=1}^N$ we look on ARCH as a two level approximation process. We shall suppose that the representation

$$y_t = f(t) + u_t \tag{2}$$

holds, where $\{u_t\}$ are independent random variables normally distributed with zero mean and variance s_t^2. The first level is to find the deterministic part $f(t)$. For this purpose the method of ordinary least squares (OLS) is often used. It is very convenient when $f(t)$ is searched as a linear combination of appropriate terms. For example by algebraic polynomials:

$$y_t = a_0 + a_1 t + \mathbf{L} + a_p t^p + u_t \tag{3}$$

or by AR(p) model:

$$y_t = a_0 + a_1 y_{t-1} + \mathbf{L} + a_p y_{t-p} + u_t . \tag{4}$$

Other forms, often used for approximation of $f(t)$ are the spline functions:

$$y_t = a_0 + a_1 t + \mathbf{L} + a_p t^p + \sum_{i=1}^{m} c_i (t - t_i)_+^p + u_t \tag{5}$$

and the exponential polynomials:

$$y_t = a_0 + a_1 e^{z_1 t} + \mathbf{L} + a_p e^{z_p t} + u_t . \tag{6}$$

Recall that $x_+^p = x^p$ if $x > 0$ and $x_+^p = 0$ otherwise. It is clear that the formula (5) is a generalization of (3) and (6) is a generalization of (4) in the case $u_t = 0$. However, the more complicated is a formula, the more difficult is the problem of the optimal choice of the parameters. The knots $\{t_i\}$ in (5) one can pick up from the graph of y_t, as a not very large number of points, where the process evidently changes it's character. Yet the method of ARCH may help for this choice as well. Next, for the exponents $\{e^{z_i}\}$ in (6) one can choose the roots of characteristic polynomial corresponding to the AR(p) model.

Using OLS estimator, one certainly obtain a residual $\{\hat{u}_t\}$ with zero mean, while for checking the hypotheses for independence and normality of $\{u_t\}$ there are various tests (see [2]). Every such test can be used as a criterion if the deterministic part is found well.

As criterions for comparing a method M with other methods we shall use:

1) Adjusted square mean $s_{adj} := s(\hat{u}_t) \sqrt{N/(N-k)}$, where k is the number of the degrees of freedom of the method;

2) Autocorrelation coefficients $r_k(\hat{u}_t)$, $k = 1,2,3$;

3) Extrapolation error
$$s_{ext} := s\left(y_{i+1} - M(y_{i+1} \mid \{y_1,..., y_i\}, \ i = m,..., N-1 \right).$$

That is, s_{ext} is the square mean of the errors of one step prognosis of the method applied over parts of the data. (Clearly, the calculation of s_{ext} needs a number of degrees of freedom, not exceeding m. In the examples we set $m = \lfloor N/2 \rfloor$.)

Note that, under the assumptions that y_t is CLRM (classical linear regressive model) and that the method recovers $f(t)$ with an error of order s_t, then the variances of s_{adj}, $r_k(\hat{u}_t)$, s_{ext} are $O(\frac{s_{adj}^2}{N-k})$, $O(\frac{1}{N-k})$, $O(\frac{s_{ext}^2}{N-m})$, respectively.

For the second level, we suppose that the deterministic part $f(t)$ is subtracted and we look for the variance s_t^2 of u_t. Formally, we can apply to u_t^2 the same methods as at the first part, but the interpretation now is different. If we choose the AR(p) model for u_t^2, we arrive at ARCH(p). This method is a particular case of the GARCH(p,q) (Generalized ARCH) model, see Bollerslev [1], for which it is set

$$s_t^2 = a_0 + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + ... + a_p u_{t-p}^2 + b_1 s_{t-1}^2 + b_2 s_{t-2}^2 + ... + b_q s_{t-q}^2.$$

The GARCH model is more realistic than ARCH, but the estimation of the parameters turns out to be a complicated nonlinear process.

Note that in a particular ARCH(p), as well as in a GARCH(p,q) estimation it can happen that some estimated value $\hat{s}_t^2$ to be negative. This is not a good sign for the applied method. To avoid such situation Silvennoinen and Teräsvirta ([5]) considered a variant of GARCH model with appropriate exponential change of the variables.

Having obtained the estimates $\hat{s}_t^2$ of the heteroscedasticity, we can return to the first level to improve the approximation of $f(t)$. For this end we may use the method of generalized least squares (GLS) with a weight function $w_t = 1/\hat{s}_t$. That is, when the variance of the stochastic term u_t is larger we take the data with smaller weight and the opposite. Also, we need the more relevant measures for the "goodness" of a method in terms of the normalized residuals $u_t^{\circ} = u_t/\hat{s}_t$).

Examples of ARCH models.

The calculations are performed by MATLAB 7.0.1.

I. ARCH model of interest

We consider the basis interest rate of BNB on monthly basis for the period 01.2008 – 09.2011.

```
Y_t=[4.68 4.75 4.83 4.85 4.93 4.96 5.06 5.25 5.23
5.38 5.72 5.77 5.17 3.92 3.49 3.53 2.27 2.35 2.24 1.71
1.57 1.46 0.61 0.55 0.39 0.24 0.18 0.17 0.18 0.20 0.18
0.17 0.17 0.17 0.17 0.18 0.18 0.19 0.18 0.19 0.21 0.22
0.17 0.18 0.18];
```

1) *Estimation of $f(t)$ by polynomials*

In order to choose the optimal degree p of the polynomial, consider the table

p	σ(u)	σ_adj	r_1(u)	r_2(u)	r_3(u)	σ_ext
0	2.1562	2.1562	0.9889	0.9721	0.9521	2.6828
1	0.9348	0.9453	0.9404	0.8508	0.7470	1.0408
2	0.8050	0.8235	0.9170	0.7885	0.6395	1.2170
3	0.5760	0.5962	0.8574	0.6735	0.4973	0.8580
4	0.3577	0.3747	0.6528	0.2943	0.0628	0.5758
5	0.3504	0.3716	0.6513	0.3029	0.0860	0.6187
6	0.2582	0.2773	0.4618	0.0482	-0.0827	0.7199
7	0.2385	0.2595	0.3623	-0.0980	-0.1848	0.7982
8	0.2362	0.2604	0.3683	-0.0817	-0.1589	1.0573
9	0.2230	0.2493	0.3333	-0.1034	-0.1335	1.2300
10	0.2086	0.2365	0.2363	-0.2160	-0.1597	1.9869
11	0.2072	0.2383	0.2145	-0.2440	-0.1758	2.6853
12	0.1933	0.2257	0.1525	-0.2811	-0.1200	4.1102

From this table we see that the choice $p = 4$ or $p = 6$ is optimal according to r_2, r_3 and σ_ext criterions. To understand the behavior of r_1 and way it is better to ignore it, let us look at the graphs of $f(t)$ and u_t.
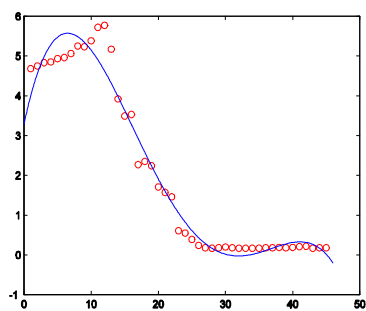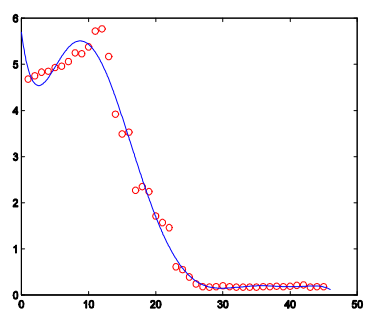
$p = 4:$ $p = 7:$ $p = 13:$
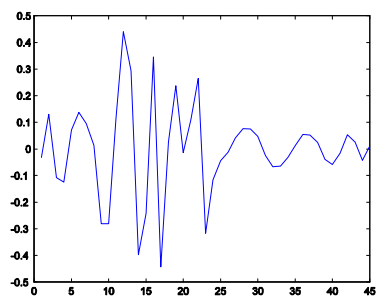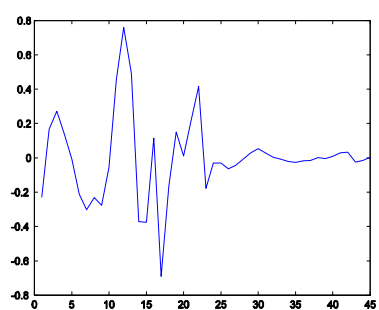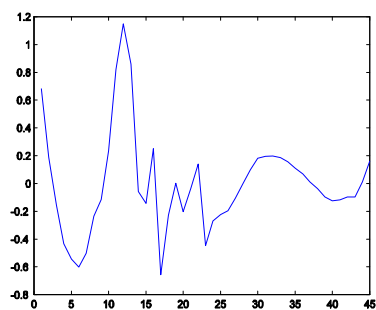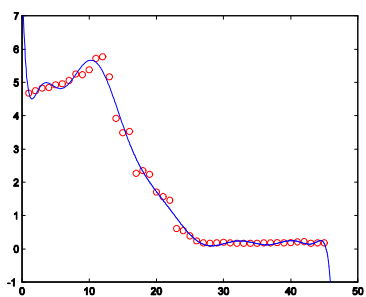
Figures 1-6

(We provide the case $p = 7$, since graphically it makes the best fit.) On the last graph we see that by chance the "large" part of u_t preferably alternate and gives "large" negative sum in r_1. This part is compensated by the series waves, which give positive sum in r_1. So, the case $p = 13$, despite the small value of r_1, do not fit to the model. As a conclusion, the "r_1 criterion" fails because of high heteroscedasticity.

Note that the first two columns also can be used for choosing the appropriate degree, but with more caution. The rule is that, around the optimal choice of the parameter, the sequence holds up its value.

Let us choose the approximation by polynomials of degree 4 and turn to the second level. We have:

```
u_t= [0.6830 0.1903 -0.1542 -0.4341 -0.5421 -0.6006
-0.5013 -0.2355 -0.1136 0.2340 0.8181 1.1496  0.8601 -
0.0583 -0.1427  0.2500 -0.6562 -0.2270  0.0026 -0.2019
-0.0344   0.1417  -0.4464  -0.2709  -0.2235  -0.1954  -
0.1071   0.0014  0.1006  0.1816  0.1962  0.1965  0.1853
0.1559   0.1123  0.0690  0.0109 -0.0364 -0.0967 -0.1233
-0.1190 -0.0959 -0.0958  0.0105  0.1623]
```

Assuming that $u_t \sim N(0, s_t^2)$, for s_t^2 we obtain the estimates:

ARCH(0): $s_t^2 = a_0$;

$$\hat{a}_0 = 0.1280;$$

$s_{adj} = 0.2521; r_1 = 0.5793, \quad r_2 = 0.0454, \quad r_3 = -0.1150, \quad s_{ext} = 0.1504.$
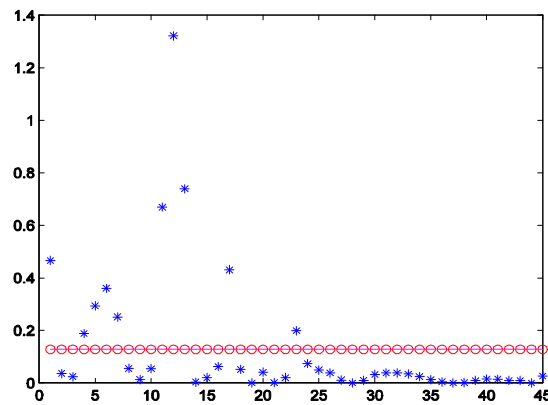
Figure 7

ARCH(1): $s_t^2 = a_0 + a_1 u_{t-1}^2$

$\hat{a}_0 = 0.0463$, $\hat{a}_1 = 0.5681$;

$s_{adj} = 0.2058$, $r_1 = 0.3052$, $r_2 = -0.3100$, $r_3 = -0.3560$, $s_{ext} = 0.0683$.
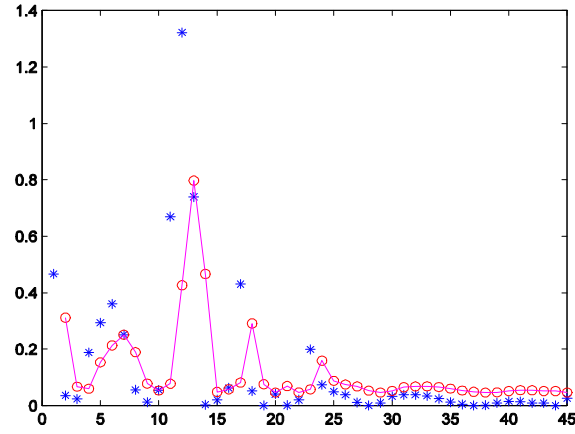
Figure 8

ARCH(2): $s_t^2 = a_0 + a_1 u_{t-1}^2 + a_2 u_{t-2}^2$

$\hat{a}_0 = 0.0757$, $\hat{a}_1 = 0.8662$, $\hat{a}_2 = -0.4466$;

$s_{adj} = 0.1824$, $r_1 = 0.1331$, $r_2 = -0.0218$, $r_3 = 0.0655$, $s_{ext} = 0.1111$.

Figure 9

51
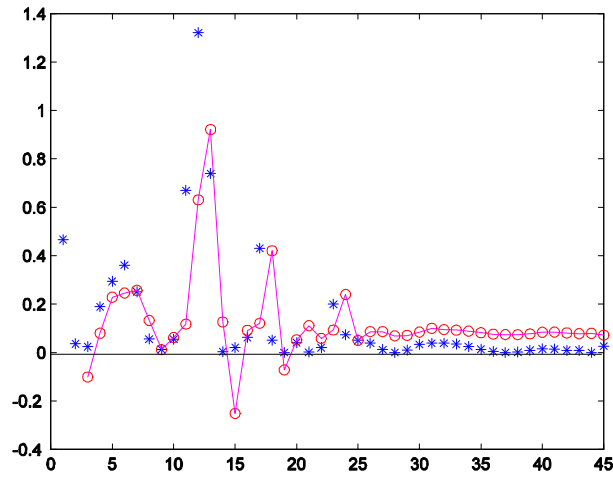
ARCH(3): $s_t^2 = a_0 + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + a_3 u_{t-3}^2$

$\hat{a}_0 = 0.0539$, $\hat{a}_1 = 0.9988$, $\hat{a}_2 = -0.6764$, $\hat{a}_3 = 0.2285$;

$s_{adj} = 0.1805$, $r_1 = -0.0428$, $r_2 = 0.0359$, $r_3 = 0.0134$, $s_{ext} = 0.0936$.

We see from these results that the best fit is given by the estimation ARCH(1). For $p \geq 2$ the presence of negative suggestions for s_t^2 make the method ARCH(p) not usable in these cases. Another observation is that the criterions by r_1, r_2 and r_3 do not work. Their values are rather small than the expected statistical error ($s(r_k) \approx 0.15$), and the effect increases with p. Probably, the reason is the close connection between the method of ARCH and autocorrelation.

2) *Estimation of $f(t)$ using AR(p)*

To choose an appropriate value of p in AR(p), let us consider the table

p	σ(u)	σ_adj	r₁(|u|)	r₂(|u|)	r₃(|u|)	σ_ext
0	2.1562	2.1806	0.9270	0.8109	0.6626	2.7201
1	0.3193	0.3268	0.1495	0.0824	0.6975	0.2039
2	0.3122	0.3237	0.2212	0.0153	0.6271	0.1939
3	0.3133	0.3294	0.2488	0.0585	0.6538	0.1955
4	0.2867	0.3060	0.4132	0.3066	0.5032	0.1722
5	0.2876	0.3119	0.3875	0.3311	0.5510	0.1766
6	0.2900	0.3201	0.3944	0.3635	0.5530	0.2076
7	0.2887	.03249	0.3978	0.3680	0.6145	0.2150
8	0.2853	0.3279	0.4040	0.3980	0.5925	2.0944
9	0.2796	0.3290	0.4302	0.4696	0.5866	2.9592
10	0.2715	0.3280	0.3870	0.4814	0.5713	3.6184

From the table and the graphs it is seen that the optimal choice is $p = 4$ and partially the choice $p = 2$. The other graphs for $p > 4$ (not given here) are very similar to that for $p = 4$, but starting from larger indexes.

p=2:   p=4:
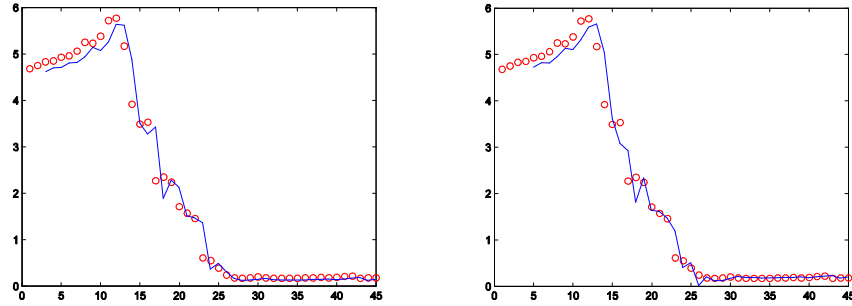
Figures 10-11

AR(2): $y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + u_t$, $\quad \hat{a}_0 = -0.0352$, $\hat{a}_1 = 1.2090$, $\hat{a}_2 = -0.2315$.

AR(4): $y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + a_4 y_{t-4} + u_t$,

$\hat{a}_0 = 0.0156$, $\hat{a}_1 = 1.1418$, $\hat{a}_2 = -0.1919$, $\hat{a}_3 = 0.4294$,

$\hat{a}_4 = -0.4134$.

Choosing the second model, for the error term we obtain:

```
    uₜ=[ 0  0  0  0   0.1983   0.1354   0.2410   0.2965
0.0989   0.2777   0.4023   0.1800 -0.4845 -1.1238 -0.1226
0.4469 -0.6526   0.5417 -0.0964   0.0722 -0.0390 -0.0105
-0.5797   0.1507 -0.1145   0.2262 -0.0188   0.0548
0.0630   0.0334 -0.0280 -0.0198 -0.0166 -0.0017 -0.0057
0.0002 -0.0112   0.0007 -0.0250   0.0024   0.0048   0.0023
-0.0637   0.0009 -0.0162].
```

Note that the unknown first four terms are set to 0, but in the calculations we use the vector $u_{t'}$ with dropped zero sequence.

The results for ARCH(p) are given at the table below. (U is the residual.)

p	$\sigma(U)$	σ_{adj}	$r_1(|U|)$	$r_2(|U|)$	$r_3(|U|)$	σ_{ext}
0	0.2120	0.2147	0.0460	-0.0273	0.1750	0.1149
1	0.2118	0.2173	0.2183	-0.0397	0.1622	0.1084
2	0.2140	0.2227	0.1887	-0.0139	0.1671	0.1144

3	0.2067	0.2186	0.1811	-0.0338	-0.0329	0.0949
4	0.2085	0.2243	0.1751	-0.0272	-0.0392	0.0940
5	0.2106	0.2308	0.1653	-0.0489	-0.0411	0.1140
6	0.2134	0.2387	0.1568	-0.0565	-0.0356	0.1375

The table does not show some decisive advantage for these p and the graphs for p=3,…,6 are very similar. However, for p>6, both the numbers and graphs show that the approximation is not acceptable.

ARCH(3): $s_t^2 = a_0 + a_1 u_{t-1}^2 + a_2 u_{t-2}^2 + a_3 u_{t-3}^2$

$\hat{a}_0 = 0.0464$, $\hat{a}_1 = 0.1341$, $\hat{a}_2 = 0.0094$, $\hat{a}_3 = 0.3017$.
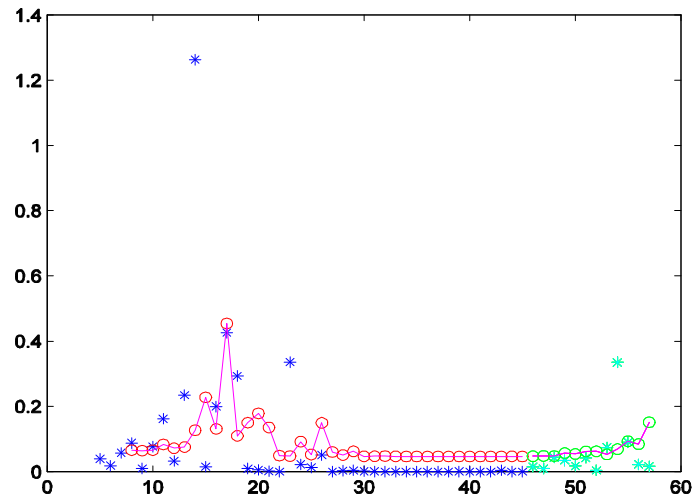
Figure 12

Figure 12 shows the current estimation of the volatility using ARCH(3) and a simulation for a year ahead.

II. ARCH model of inflation

We consider the US monthly inflation percent for the period 01.2007 – 03.2012. The data was taken from the website InflationData.Com .

```
Y_t=[2.08 2.42 2.78 2.57 2.69 2.69 2.36 1.97 2.76
3.54 4.31 4.08 4.28 4.03 3.98 3.94 4.18 5.02 5.60 5.37
```

```
4.94 3.66 1.07 0.09 0.03 0.24 -0.38 -0.74 -1.28 -1.43
-2.10 -1.48 -1.29 -0.18 1.84 2.72 2.63 2.14 2.31 2.24
2.02 1.05 1.24 1.15 1.14 1.17 1.14 1.50 1.63 2.11 2.68
3.16 3.57 3.56 3.63 3.77 3.87 3.53 3.39 2.96 2.93 2.87
2.65];
```

```
Annual 2007-2011: 2.85 3.85 -0.34 1.64 3.16
```

The table for the criterions that we use is as follows:

p	σ(u)	σ$_{adj}$	r$_1$(|u|)	r$_2$(|u|)	r$_3$(|u|)	σ$_{ext}$
0	1.7725	1.7868	0.8969	0.7261	0.5434	1.3613
1	0.6109	0.6209	0.4204	-0.0892	-0.1343	0.5571
2	0.4906	0.5031	0.2010	0.0441	-0.1109	0.4478
3	0.4940	0.5114	0.1985	0.0296	-0.1369	0.4695
4	0.4952	0.5176	0.1872	0.0331	-0.1127	0.4784
5	0.4910	0.5186	0.2466	0.0713	-0.1495	0.4853
6	0.4937	0.5271	0.2620	0.0205	-0.1103	0.5057

Clearly, the best choice is $p = 2$ and we have:

AR(2): $y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + u_t$, $\hat{a}_0 = 0.2136$, $\hat{a}_1 = 1.5054$, $\hat{a}_2 = -0.6018$.

Estimating the volatility by ARCH(p), we get that the best assumption is $p = 0$, that is homoscedasticity with $s_t^2 = b_0$, $\hat{b}_0 = 0.2407$.

By a Monte-Karlo experiment with 1000 runs of ARCH(0) we find a prognosis for annual inflation for 2012 with 95% confidence interval: | Infl - 2.36 | < 0.88.
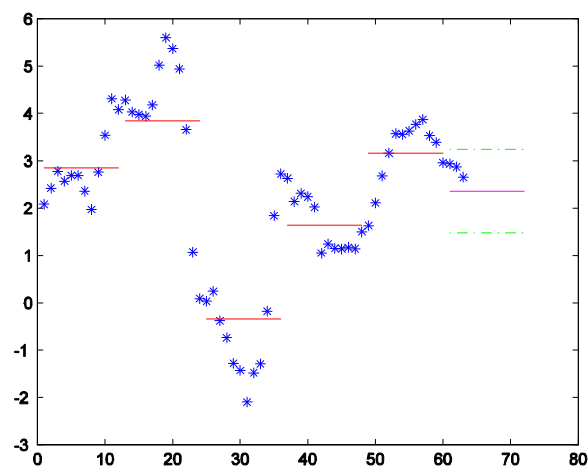
Figure 13

References

1. Bollerslev, T. "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31 (1986), 307–327.
2. Damodar, N. Gujarati, Basic econometrics, 1995.
3. Engle, R. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50 (1982), 987–1008.
4. Engle, R. F. and Rangel, J. G. The Spline GARCH Model for Unconditional Volatility and its Global Macroeconomic Causes, Czech National Bank, 2005.
5. Silvennoinen, A., Teräsvirta, T. Modelling Multivariate Autoregressive Conditional Heteroscedasticity with the Double Smooth Transition Conditional Correlation GARCH Model, SSE/EFI Working Paper Series in Economics and Finance No. 652 (2007).

Authors:
assistant prof. **Savka Kostadinova**
Sofia University St. Kl. Ohridski, Faculty of Matematics and Informatics,
0899152273, savka@fmi.uni-sofia.bg

associate prof. **Nikola Naidenov**
Sofia University St. Kl. Ohridski, Faculty of Matematics and Informatics,
nikola@fmi.uni-sofia.bg

AUTOMATION OF THE PROCESSES OF MAKING THE CHIEF BUDGET OF A COMPANY

Ilia Lichev, Mariya Deshkova

Abstract: *This report explores the problem of automation of the production budget of a company, and it is based on the requirements of the accounting at home and is made through software package ACCESS. For this purpose, the database is formed in tables, which are needed for complex computational processes in order to obtain the main budget tables and references, as well as the budget tables themselves.*

Key words: *database, chief budget, budget tables*

This report attempts to propose a technology for preparing the main budget of the enterprise, taking advantage of software package MS ACCESS.

The automation of processes in the preparation of the main enterprise budget includes two phases:

1. Description of organizational and economic nature of the task

The budget process is a set of actions determining the future behavior of the individual parts of the system, reflected in the individual budgets. The budget process transforms the long-term plan into an annual plan, divided into month quarters. The main budget of the organization is prepared on the basis of individual budgets.

The main budget is a comprehensive budget that expresses the organization's overall plan for one year or less periods (quarters, months). It enables to fulfill the following:

– promotes interaction between different activities of the organization

– combines the individual budgets of the separate parts

– combines individual budgets in a harmony for the company

The main budget contains two types of budgets:

– Operating budget

– Financial budget

The operating budget reflects the results of the companies activities during the budget period. It includes plans to implement various activities in different parts of the system - production, sales, personnel, and expectations of such important areas as profit, expenses, income, etc. In the operating budget reflects the activities of the individual parts of the system as a whole and interdependence.

The financial budget plan includes revenue and expenditure of cash (cash flow) plan for kept profits and preparation of project balance of the enterprise.

Preparation of the main Budget of the company includes the following phases:

- Analysis of the balance of the company during a previous period
- Preparation of budget tables, namely:

– budget table for sales per months of the quarter and areas for individual products.

– budget table of the amounts to be paid by purchasers in months of the quarter.

– budget table for the specified products to produce monthly and quarterly total for the period. Data is taken from individual Budget of operations.

– budget table for the expected production of goods per months of the quarter. Information for the preparation of this table are also taken by individual budget operations.

– budget table for direct labor costs for the preparation of individual products. Data is taken from the individual budget of the staff.

– budget table for the amount of overhead costs of each department and all of the company. Give the controllable and uncontrollable costs for each element. For this purpose we used the processed before that standards for overhead costs to individual departments.

– budget table for production costs.

– budget table for sales and administrative costs.

– budget table for planned costs of individual units, departments, divisions etc.

– budget table for the enterprise divided into elements.

– budget table for the amount necessary to purchase materials for a total enterprise

– budget table for the amounts paid for the materials will be purchased during the planning period.

– budget table for the planned conversion cost.

– budget table for the receipt and disbursement of money. In summary it reflects all areas of revenue and expenditure for the next period.

– budget table for production value of goods sold.

– budget table for the planned income of the enterprise.

– budget table for retained earnings.

2. Projecting manufacturing phase

Based on the above a processing stage was designed by developing a database to prepare the main budget of the company through ACCESS. Processing phase involves the following procedures:

a) Menu Design

The task starts with a macro named Autoexec as soon as you activate the file **COSTMAN.MDB** in an environment of ACCESS, the macro opens the title screen with the name of the problem and systematically in a horizontal menu bar containing the main activities in the scheme. In choosing an activity-menu option opens a drop-down menu from which you / by clicking with the mouse or by the underlined letter in the name of the option from this menu / choose another activity for building the database, update the database and display the relevant budget table from the database.

Options on the horizontal menu bar are: **Create a table, Update tables, Budget tables** and **Exit** presented in Figure 1.
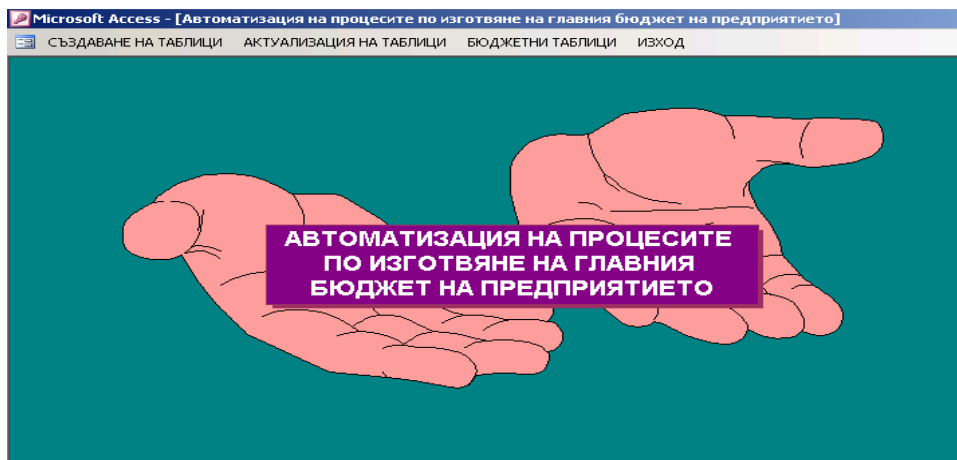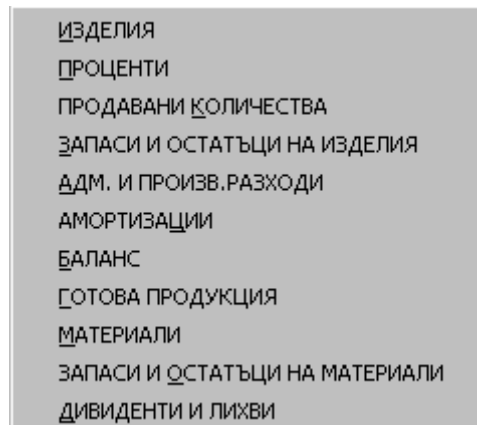
Figure 1

b) Creating Tables

Create tables option allows to form the database by creating separate files / tables /.This is shown in the submenu, including 11 tables, depicted in Figure 2.

Creating tables:

ИЗДЕЛИЯ
ПРОЦЕНТИ
ПРОДАВАНИ КОЛИЧЕСТВА
ЗАПАСИ И ОСТАТЪЦИ НА ИЗДЕЛИЯ
АДМ. И ПРОИЗВ.РАЗХОДИ
АМОРТИЗАЦИИ
БАЛАНС
ГОТОВА ПРОДУКЦИЯ
МАТЕРИАЛИ
ЗАПАСИ И ОСТАТЪЦИ НА МАТЕРИАЛИ
ДИВИДЕНТИ И ЛИХВИ

Figure 2

Table **products** to record information on products, including their code, name, measure, and price.

Table **Percentages** contains monthly rates paid in cash and loan value of the products.

Sold volumes that contains data about the code of the products and the quantities planned for sale in months.

Stocks of products **and residues,** which contain information about the code of the device and planned for it reserves and scrap supplies.

Administrative and production costs, which contain information about variable and fixed costs per month.

Depreciation, which contains data on production and administrative expenses by month.

Balance which contains data for code name of the item in the balance and balance remaining on it.

Finished goods containing data on finished goods /received, sold and balance/ articles per months.

Materials Table to record information on materials, including their code, name, measure, and price.

Inventories of materials and residues, which contains data for the code of the material and planned for it stocks and residues

Dividends and interest, which contains data on interest paid and dividends on a quarterly basis.

The structure of these tables is constructed through the site Table of ACCESS. For the completion of these data, tables, and   the

construction of the actual database we are using input forms built through the site FORMS of ACCESS and more accurately by AutoForm or Form Wizard.

Here are the contents of the entry forms to fill some of the tables - from fig. 3 to fig. 6:

Figure 3

Figure 4

Figure 5

Figure 6

Entry forms for making the data in the tables are constructed in the usual way, by allowing the navigation buttons at the bottom of the form to move between records, following the order of their introduction. As we position on a record we can look at it, rectify or delete using the Delete button on the standard toolbar buttons ACCESS. As we close the form by clicking Close, the base table that is selected by the form, the change takes place. This is the easiest and most affordable way to work with forms in ACCESS.

In the project however are used specially designed forms to fill in the data tables in which navigation buttons are missing. To save the entered data is used the built form button "Save" and to exit the form without saving - click "Cancel". Thanks to the buttons attached to macros are performed the above actions.

c) Update tables

The second option **Update tables** is using other features of ACCESS, namely working with Macros object and parametric forms. In these new forms have built-in buttons for Edit, Cancel and delete data from base tables. Through macros attached to buttons are performed appropriate actions. In a particular field in the parametric forms constructed as PopUp forms is entered information for the searched record in the database /normally this is the key of the record or part of it when it is composed of more than one field/. If you choose "Edit" button, a macro is responsible for finding the record, and its positioning on the form for editing. The "Cancel" closes the form without parameter to search the record and the button "Delete" deletes this record from the table in a database that meets the specified code in the box. Example of

62

a parametric form for updating the table "Products" is given in Fig.7. For the remaining tables are created similar forms.
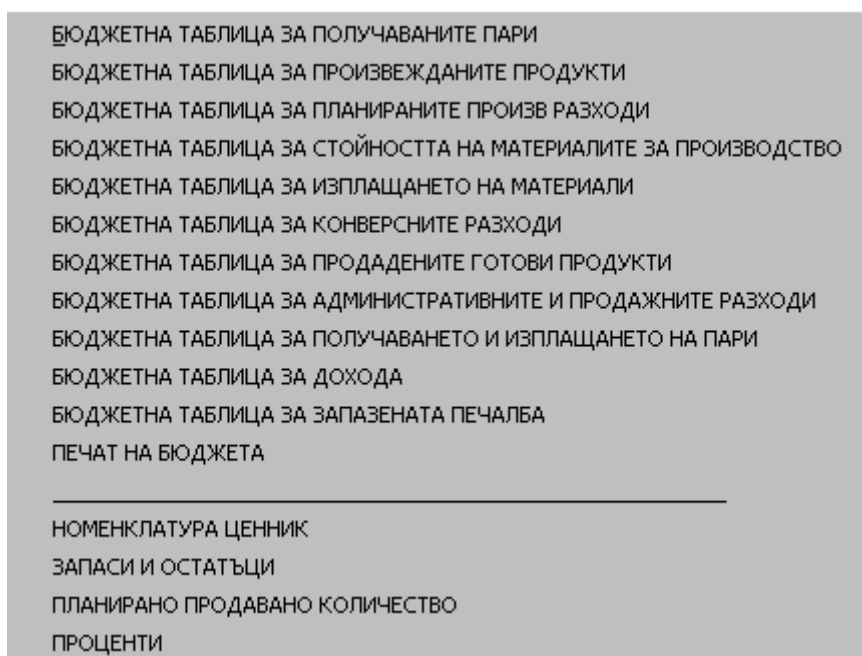
Figure 7

3. Printing **Reports**

The third option **Budget tables** includes the following items in the submenu shown in figure 8:

Figure 8

63

In those reports from the second group in the submenu we can print data from a corresponding code tables. As an example we show that nomenclature of articles in figure 9.

изделия

код	наименование
001	бойлер
002	пералня
003	хладилник
004	фритюрник
005	печка

Figure 9

To receive the budget tables given in the third section of the menu we are using a complex system of over 65 queries, built on the tables, and over than 20 macros.

Queries are other sites in ACCESS, which are based on basic spreadsheets or other applications using their various fields / elements /, and they can use certain parameters for extraction /filtering/ records to perform various computational processes between fields, to sort and/or group records from certain tables for the needs of various source documents /reports, reports/. Macros are programs that combine the use of macros and operators of SQL

The creation of these documents are made not on the tables themselves, but on requests made by Query object to ACCESS. These requests include procedures as linking records from different tables, sorting by certain signs or certain parameters extraction, grouping and summarizing data, and not least in importance performing various computing processes, the results of which are kept in so-called computing fields. For the request is typical that the data is not permanently stored in disk memory, but is available in the execution of the application. Only when using Make Table Query actually a table is created on disk when necessary. In our base such applications are designed to store data for the quarter as it works on two levels - by months and quarters. By request to delete – Delete Query, adding – Append Query and update – Update Query can be performed the above

64

actions and the order of its usage is given in the macros in order to receive the result – budget tables.

For the preparation of screen "Budget Table for received money" a system is used that have many applications, the latter of which is shown in Figure 10 and macro, presented in Figure 11.

Figure 10

This field Total is given by the formula:

ОБЩО: [SumOfИЗПЛ-СУМИ]+[SumOfТЕК ИЗПЛ КРЕДИТ]+[ОТ ПРЕДХ МЕС]

Figure 11

Due to the limited volume of the report the requests will not be described further, as well as the macros.

Further on we will describe only part of the contents of the budgetary tables generated by the developed system and via the above procedure /figure 12 – figure 26/.

БЮДЖЕТНА ТАБЛИЦА ЗА ПОЛУЧАВАНИТЕ ПАРИ

ЗА ТРИМЕСЕЧИЕ 1

МЕСЕЦ	СТОЙНОСТ	ИЗПЛ.СУМИ	КРЕДИТ	ТЕК.ИЗПЛ.КРЕДИТ	ОТ ПРЕДХ.МЕС	ОБЩО
1	250 000.00 лв	62 500.00 лв	187 500.00 лв	93 750.00 лв	75 000.00 лв	231 250.00 лв
2	500 000.00 лв	125 000.00 лв	375 000.00 лв	187 500.00 лв	93 750.00 лв	406 250.00 лв
3	600 000.00 лв	150 000.00 лв	450 000.00 лв	225 000.00 лв	187 500.00 лв	562 500.00 лв
ОБЩО:	1 350 000.00 лв	337 500.00 лв	1 012 500.00 лв	506 250.00 лв	356 250.00 лв	1 200 000.00 лв

Figure 12

БЮДЖЕТНА ТАБЛИЦА ЗА ПЛАНИРАНИТЕ ПРОИЗВОДСТВЕНИ РАЗХОДИ

ЗА ТРИМЕСЕЧИЕ 1

месец	пр. разходи материали	пр. разходи на труд	пром. реж. разходи	постоянни реж. разх	общо произв. разходи
1	75 000.00 лв	90 000.00 лв	45 000.00 лв	15 000.00 лв	225 000.00 лв
2	110 000.00 лв	132 000.00 лв	66 000.00 лв	15 000.00 лв	323 000.00 лв
3	105 000.00 лв	126 000.00 лв	63 000.00 лв	15 000.00 лв	309 000.00 лв
ОБЩО:	290 000.00 лв	348 000.00 лв	174 000.00 лв	45 000.00 лв	857 000.00 лв

Figure 13

Бюджетна таблица за произвежданите продукти

ЗА ТРИМЕСЕЧИЕ 1

ИЗДЕЛИЕ	МЕСЕЦ	ПРОД.КОЛ-ВО	ЗАПАСИ	ОСТАТЪЦИ	ПЛАН-ПРОИЗВОДСТВО
БОЙЛЕР					
	ЯНУАРИ	10000	10000	5000	15000
	ФЕВРУАРИ	20000	12000	10000	22000
	МАРТ	24000	9000	12000	21000
Сума за 'изделие' = БОЙЛЕР (3 дет. записа)		54000	31000	27000	58000
ОБЩО ЗА ТРИМЕСЕЧИЕТО:		54000	31000	27000	58000

Figure 14

БЮДЖЕТНА ТАБЛИЦА ЗА СТОЙНОСТТА НА НЕОБХОДИМИТЕ ЗА ПОКУПКА ОСНОВНИ И СПОМАГАТЕЛНИ МАТЕРИАЛИ ЗА ПРОИЗВОДСТВО ЗА ТРИМЕСЕЧИЕ [1]

месец	запаси	пр. разходи матер.	общо нужди	остатъци	за покупка
1	10 000,00 лв	75 000,00 лв	85 000,00 лв	8 000,00 лв	77 000,00 лв
2	10 000,00 лв	110 000,00 лв	120 000,00 лв	10 000,00 лв	110 000,00 лв
3	10 000,00 лв	105 000,00 лв	115 000,00 лв	10 000,00 лв	105 000,00 лв
ОБЩО:	30 000,00 лв	290 000,00 лв	320 000,00 лв	28 000,00 лв	292 000,00 лв

Figure 15

БЮДЖЕТНА ТАБЛИЦА ЗА ИЗПЛАЩАНЕТО НА МАТЕРИАЛИ ЗА ТРИМЕСЕЧИЕ [1]

месец	стойност на необх. мат-ли за покупка	суми за изплащане
1	77 000.00 лв	46 000.00 лв
2	110 000.00 лв	77 000.00 лв
3	105 000.00 лв	110 000.00 лв
ОБЩО:	292 000.00 лв	233 000.00 лв

Figure 16

БЮДЖЕТНА ТАБЛИЦА ЗА КОНВЕРСНИТЕ РАЗХОДИ ЗА ТРИМЕСЕЧИЕ [1]

МЕСЕЦ	ПР. РАЗХОДИ ТРУД	ПРОМ. РЕЖИЙНИ	ПОСТ. РАЗХОДИ	ОБЩО РАЗХОДИ
1	90 000.00 лв	45 000.00 лв	13 000.00 лв	148 000.00 лв
2	132 000.00 лв	66 000.00 лв	13 000.00 лв	211 000.00 лв
3	126 000.00 лв	63 000.00 лв	13 000.00 лв	202 000.00 лв
ОБЩО:	348 000.00 лв	174 000.00 лв	39 000.00 лв	561 000.00 лв

Figure 17

Figure 18

Figure 19

Figure 20

БЮДЖЕТНА ТАБЛИЦА ЗА ДОХОДА

ЗА ТРИМЕСЕЧИЕ: 1

ПРОДАЖБИ	1 350 000.00 лв
СТОЙНОСТ НА ПРОДАДЕНИТЕ ПРОДУКТИ	799 017.24 лв
БРУТНА ПЕЧАЛБА	550 982.76 лв
АДМИН. И ПРОДАЖНИ РАЗХОДИ	132 000.00 лв
ОПЕРАЦИОНЕН ДОХОД	418 982.76 лв
ЛИХВИ	1 097.00 лв
ЧИСТ ДОХОД	417 885.76 лв

Figure 21

БЮДЖЕТНА ТАБЛИЦА ЗА ЗАПАЗЕНАТА ПЕЧАЛБА

ЗА ТРИМЕСЕЧИЕ: 1

ЗАПАЗЕНА ПЕЧАЛБА ОТ МИНАЛАТА ГОДИНА	228 000.00 лв
ЧИСТ ДОХОД	417 885.76 лв
ДИВИДЕНТИ	80 000.00 лв
ЗАПАЗЕНА ПЕЧАЛБА	565 885.76 лв

Figure 22

НОМЕНКЛАТУРА-ЦЕННИК

наименование	шифър изделие	мярка	цена
БОЙЛЕР	01	БР	25.00 лв

Figure 23

ЗАПАСИ И ОСТАТЪЦИ

НАИМЕНОВАНИЕ НА ИЗДЕЛИЕ	МЕСЕЦ	ПЛ. ЗАПАС	ПЛ. ОСТАТЪК
БОЙЛЕР			
	ЯНУАРИ	10000	5000
	ФЕВРУАРИ	12000	10000
	МАРТ	9000	12000

Figure 24

ПЛАНИРАНО ПРОДАДЕНО КОЛИЧЕСТВО

ИЗДЕЛИЕ	МЕСЕЦ	ПЛАНИРАНО ПРОДАДЕНО КОЛИЧЕСТВО
01	1	10000
01	2	20000
01	3	24000

Figure 25

ПРОЦЕНТИ НА ПЛАЩАНЕ

ТРИМЕС.	МЕСЕЦ	НАИМЕНОВАНИЕ	% В БРОЙ	% КРЕДИТ	% ТЕК. ИЗПЛ КРЕДИТ
1	1	ЯНУАРИ	25	75	50
1	2	ФЕВРУАРИ	25	75	50
1	3	МАРТ	25	75	50
2	4	АПРИЛ	30	70	50
2	5	МАЙ	30	70	50
2	6	ЮНИ	30	70	50

Figure 26

By the last option in the horizontal bar **Exit** we are leaving the system.

70

As a result of the foregoing, we could point out that the automation of processes for the preparation of the main budget of the industrial companies is a highly complex process that requires the use of procedures enabling its rationalization and optimization. In this regard, we emphasized the potential of ACCESS software and its tools to achieve this automation.

The developed system could help to increase efficiency and improve the management system in the industrial business.

References
1. Bocheva, Penka. Development of Business Information Systems with MS ACCESS 2.0 1998
2. Lichev, Ilia. Management Accounting University Press of University of Economics Varna, 1997
3. Manuals for ACCESS 2000 and 2003

Authors:
Ilia Krumov Lichev – Professor Dr. University of economics, phone 052784223
Mariya Deshkova Dimitrova – VSU assistant, phone 052653019,
e-mail deshkova@abv.bg

HIERARCHICAL APPROACH FOR OPTIMAL ALLOCATION OF FINANCIAL RESOURCES

Krasimira Stoilova

Abstract. *The optimal resource allocation of sets of securities, available at the financial market can be done by solving multi-criterion optimization task, The maximization of portfolio return and minimization of portfolio risk is the main goal at the financial market where a set of securities has to be allocated. These contradictionary goals are formalized by solving mini-max optimization problem, which does not have a trivial solution. The investment process is presented by hierarchical approach in the paper. The decision making process is realized here by objective criteria and the subjective element, typical for the classical case of this problem, is overcome. The problems at the upper and lower level of bi-level hierarchical system are determined.*

Key words: *Financial investments, portfolio optimization, multilevel hierarchical theory*

Portfolio optimization problem

The financial resource allocation is related with investing capital in financial assets in a manner to receive later return. The optimal resource allocation in investment process is formalized by so called portfolio optimization problem [7]. This is sophisticated optimization problem because its goal is to maximize the return and at the same time to minimize the investment risk [2, 4, 5, 6]. The peculiarity of this problem is the risk, undetermined value, corresponding to uncertainty. It reflects the dynamical financial changes and their unpredictable behavior. The solution of the portfolio problem depends on the level of investor's risk in comparison with the level of portfolio return. In that manner, the relation between return and risk is the main parameter for the investor in portfolio optimization problem. The risk of financial portfolio is equivalent to the market risk which is related to the dynamical changes of the market prices of equity, foreign exchange rates, interest rates, commodity prices [3].

The portfolio theory is developed to support decision making for investment allocation of financial assets at the stock exchange [1]. The portfolio contains a set of securities. The asset is a manner for future income for the investor. The problem of portfolio optimization targets the optimal resource allocation in investment process of trading financial assets [12]. The analytical relations among the portfolio risk V_p, portfolio return E_p and the values of the investment per type of assets x_i , according to the portfolio theory, are [7]:

$$E_p = \sum_{i=1}^{n} E_i x_i = E^T x \quad V_p = \sum_{j}^{n} \sum_{i=1}^{n} x_i x_j \operatorname{cov}(i,j) = x^T \operatorname{cov}(.)x \,,$$

where

E_i - the average value of the return of asset i;

$E^T = (E_1,..., E_n)^T$ - a vector with dimension 1 x n;

$\operatorname{cov}(i,j)$ – the co-variation coefficient between the assets i and j.

The quantitative assessment of the portfolio risk is formalized by the component $V_p = x^T \operatorname{cov}(.) x$.

The quantitative evaluation of the portfolio return is formalized by $E_p = E^T x$.

The portfolio problem solutions x_i, $i=1,n$ determine the relative amounts of the investment per security i.

The co-variation matrix is calculated from previously available statistical data for the returns of assets i - $R_i^{(1)}, R_i^{2}, \mathbf{L}, R_i^{(N)}$ and j - $R_j^{(1)}, R_j^{2}, \mathbf{L}, R_j^{(N)}$ for discrete time moments (1), (2),..., (N). The co-variation matrix is a symmetrical matrix

$$\operatorname{cov}(.) = \begin{Vmatrix} \operatorname{cov}(1,1) & \operatorname{cov}(1,2) & \mathbf{L} & \operatorname{cov}(1,n) \\ \operatorname{cov}(2,1) & \operatorname{cov}(2,2) & \mathbf{L} & \operatorname{cov}(2,n) \\ \mathbf{M} & & & \\ \operatorname{cov}(n,1) & \operatorname{cov}(n,2) & \mathbf{L} & \operatorname{cov}(n,n) \end{Vmatrix}_{n x n} ,$$

which components are calculated from the expression

$$\operatorname{cov}(i,j) = \frac{1}{N}\Big[(R_i^{(1)} - E_i)(R_j^{(1)} - E_j) + (R_i^{(2)} - E_i)(R_j^{(2)} - E_j) + \mathbf{L} + (R_i^{(N)} - E_i)(R_j^{(N)} - E_j)\Big]$$

where

$$E_i = \frac{1}{N}\Big[R_i^{(1)} + R_i^{(2)} + \mathbf{L} + R_i^{(N)}\Big]$$

$$E_j = \frac{1}{N}\Big[R_j^{(1)} + R_j^{(2)} + \mathbf{L} + R_j^{(N)}\Big]$$

are the average profits of assets i and j for the period $T = [1,2,....,N]$.

The portfolio theory defines the so called "standard" optimization problem [7] :

$$\min_{x}[\frac{1}{2}x^T \operatorname{cov}(.)x - sE^T x]$$, (1)

$$x^T.\mathbf{1} = 1$$, (2)

where

cov(.) – a symmetric positively defined square matrix n x n ,

E – (n x 1) vector of the average profits of the assets for the period of time $T = [1,2,....,N]$;

$$1 = |1\mathbf{L}\ 1|^T$$ - unity vector, n × 1;

s - a parameter of the investor's preferences to undertake risk in the investment process.

The constraint (2) means that the investment per each asset has to be done for the all investment resources (presented like 1):

$$x_1 + x_2 + \mathbf{L} + x_n = 1$$.

The solutions of the portfolio problem x_i , $i=1,n$ are the relative values of the investments, allocated for assets i, $i=1,n$.

The components of the goal function have the following meaning:

$$V_p = x^T \operatorname{cov}(.)x$$ - the quantitative assessment of the portfolio risk;

$$E_p = E^T x$$ - the quantitative value of the portfolio return.

The portfolio optimization problem's goal is minimization of the portfolio risk V_p and maximization of its return E_p. The parameter s is changing in the range [0, + ∞] and formalizes the investor's tendency to undertake risk. This is unknown parameter and that is why the portfolio problem is solved many times for different values of this coefficient. The values of s are determined by the investor, which cause its subjective character. s is important parameter because it influences the definition and later the solution of the portfolio optimization problem.

The financial analyzer chooses the optimal portfolio from the upper set of admissible solutions named "efficiency frontier" in the space risk-return $V_p=V_p(E_p)$. The "efficiency frontier" is not determined easy. Each point of this curve is found after solving the portfolio optimization problem

with different values of S. The "efficient frontier" is evaluated point after point according to the following iterative numerical procedure:

1. Choice of initial value of S for the investor's preferences. A good starting point is S =0. It corresponds to the case of investor who is not keen in risky decisions;

2. The portfolio problem is solved with the chosen S and the optimal solution $x(\sigma)$ is found.

3. Evaluation of portfolio risk V_p and portfolio return E_p. These values give a point into the space $V_p = V_p(E_p)$, which belongs to the efficient frontier;

4. New value $\sigma_{new} = \sigma_{old} + \Delta$ is chosen. Jump to point 2.

Problem (1) is solved with number of values of S. For each solution of the portfolio optimization problem one point of the efficient frontier is found in the space $V_p = V_p(E_p)$. A set of solutions $x(S)$ is found while the best value of S * for that investor is empirically estimated and the optimal portfolio solution $x(S$ *) is determined.

The subjective investor's influence is overcome in the presented model. The decision making process is formalized by two hierarchically interconnected optimization problems, Fig.1. The parameter S is determined analytically, i.e. there is not any subjective influence. This optimization problem can be evaluating of such S, which leads to good ratio "portfolio risk - return". At the lower hierarchical level the standard portfolio optimization problem is solved.
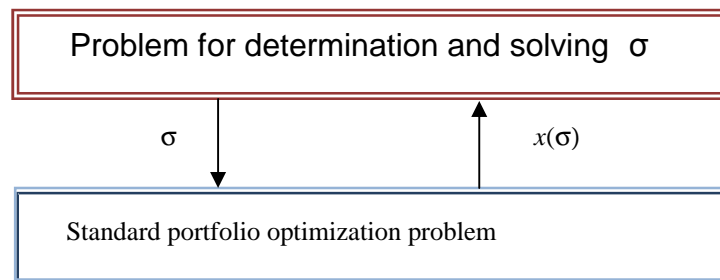
Fig. 1. Hierarchical solution of the portfolio optimization problem

Portfolio problem's solution using hierarchical approach

The hierarchical approach for solving optimization problems determines solution of optimization problem on the upper level, which solution influences the lower level optimization problems. The lower level solu-

tions define on its turn a set of parameters for the upper level problem. Because of the interrelation between the upper and lower level optimization problems by their solutions the exact form of the optimization problems cannot be defined. This complexity of the related bi-level optimization problems is solved by numerical iterative calculations of both problems (at lower and upper hierarchical level) till reaching the optimal solutions. These problems are usually nonlinear optimization problems from mathematical programming. Due to the methodological difficulties for the solution of hierarchically interconnected optimization problems, today the classical application of the portfolio theory lacks in solving bi-level optimization problems. In the paper the portfolio problem is solved by quantitative assessment of $s*$ in advance, without applying interconnected hierarchical optimization. While $s*$ is estimated intuitively or empirically by an expert for the classical case, here for the solution of bi-level portfolio problem is applied methodology, called non-iterative coordination [17, 19].

The initial problem (1) has solutions x_i, which have to be presented like analytical functions of S. The initial problem (1) is rewritten in the form

$$\min_{x}[\frac{1}{2}x^T Q x + R^T x]$$

$$Ax = C ,$$

(3)

where the correspondence between (1) and (3) is:

$$Q=\text{cov}(.) , \quad R=-S\,E , \quad A=1 , \quad C=1 .$$

If S is known, problem (1) has a solution, denoted like $x(S)$. For the case when S varies, the solution of the portfolio problem x is an inexplicit analytical function of S, or

$$x=x(S).$$

Respectively, the portfolio risk

$$V_p(S) = x^T(S)\text{cov}(.)x(S)$$

and the portfolio return

$$E_p(S) = E^T x(S)$$

are also inexplicit functions of S.

Problem (3) can be solved, applying the method of the non-iterative coordination, which gives possibility to be derived approximations of the inexplicit analytical relations of the portfolio parameters $V_p(S)$, $E_p(S)$, $x(S)$ towards the argument S for optimization problems

76

of quadratic programming. Using results from [8, 9], the analytical solution of problem (3) is

$$x^{opt} = -Q^{-1}[R - A^T(AQ^{-1}A^T)^{-1}(AQ^{-1}R + C)] . \qquad (4)$$

Using this relation, the analytical descriptions of the portfolio risk and return become

$$V_p = x^{Topt}Qx^{opt} = \left\{-\left[(C^T + R^TQ^{-1}A^T)(-AQ^{-1}A^T)^{-1}A + R^T\right]Q^{-1}\right\}Q\left\{-Q^{-1}\left[R - A^T(AQ^{-1}A^T)^{-1}(AQ^{-1}R+C)\right]\right\}$$

After transformations it follows

$$V_p = R^TQ^{-1}\left[R - A^T(AQ^{-1}A^T)^{-1}A\right]Q^{-1}R + C^T(AQ^{-1}A^T)^{-1}C .$$

The analytical relation of the portfolio return is obtained as a linear relation towards x_{opt} or

$$E_p = E^Tx = R^Tx^{opt} = R^T\left\{-Q^{-1}\left[R - A^T(AQ^{-1}A^T)^{-1}(AQ^{-1}R+C)\right]\right\} = -R^TQ^{-1}[R - A^T(AQ^{-1}A^T)^{-1}AQ^{-1}R] + R^TQ^{-1}A^T(AQ^{-1}A^T)^{-1}C$$

Finally, it holds

$$V_p = R^TQ^{-1}\left[R - A^T(AQ^{-1}A^T)^{-1}A\right]Q^{-1}R + C^T(AQ^{-1}A^T)^{-1}C \qquad (5)$$

$$E_p = -R^TQ^{-1}[R - A^T(AQ^{-1}A^T)^{-1}AQ^{-1}R] + R^TQ^{-1}A^T(AQ^{-1}A^T)^{-1}C . \qquad (6)$$

Relations (5) and (6) can be expressed in terms of the initial portfolio problem (1). Thus, explicit analytical relations for the portfolio risk V_p, portfolio return E_p and the optimal solution of the portfolio problem x_{opt} are derived towards the coefficient of the investor's risk preference s. For the current problem (1), taking into account the correspondence between problems (1) and (3), it follows

$$x^{opt}(s) = Q^{-1}\left\{\left[E - A^T(AQ^{-1}A^T)^{-1}AQ^{-1}E\right]s + A^T(AQ^{-1}A^T)^{-1}C\right\} \qquad (7)$$

$$V_p(s) = E^TQ^{-1}[E - A^T(AQ^{-1}A^T)^{-1}AQ^{-1}E]s^2 + C^T(AQ^{-1}A^T)^{-1}C \qquad (8)$$

$$E_p(s) = E^Tx^{opt}(s) = E^TQ^{-1}\left\{\left[E - A^T(AQ^{-1}A^T)^{-1}AQ^{-1}E\right]s + A^T(AQ^{-1}A^T)^{-1}C\right\} \qquad (9)$$

The following notations are used to simplify the expressions

$$a = E^TQ^{-1}\left[E - A^T(AQ^{-1}A^T)^{-1}AQ^{-1}E\right] \qquad (10)$$

$$b = C^T (AQ^{-1}A^T)^{-1}C$$
$$g = E^T Q^{-1} A^T (AQ^{-1}A^T)^{-1}C \, ,$$

where the parameters α, β and γ are scalars. Respectively, using (10), relations (8) and (9) become

$$V_p(s) = as^2 + b \, ,$$
$$E_p(s) = as + g \, . \tag{11}$$

The derived relations (7), (8), (9) and (10) describe in analytically explicit form the functional relations between the portfolio parameters for risk, return and optimal solution towards the coefficient of the investor's preferences to risk s. Hence, the solution of the portfolio problem (1) is calculated using relations (7)-(9) without implementation of optimization algorithms for the solution of the low level optimization problem. This considerably speeds up the problem solution of (1). Hence, the portfolio optimization problem can be solved in real time, with lack of iterative calculations, which benefits the decision making in fast dynamical environment of stock exchange.

On the upper optimization level it is necessary to be evaluated σ - the parameter of investor's preferences, under which the better (minimal) value of the relation Risk/Return for the optimal portfolio is achieved. This relation is known as Sharp ratio. The formalization problem for evaluation of σ is

$$\min_{s \geq 0} \left\{ \frac{Risk(s)}{Portfolio\_return(s)} = \frac{V_p(s)}{E_p(s)} \right\}.$$

According to relation (11) the analytical form of the problem is

$$\min_{s} \left\{ \frac{V_p(s)}{E_p(s)} = \frac{as^2 + b}{as + g} = h(s) \right\} \, . \tag{12}$$

This problem evaluates the parameter of the investor's preferences σ according to objective considerations. Thus, the portfolio optimization problem is stated like bi-level optimization procedure, Fig. 1. The advantage for the evaluation of σ comes from the fact that the estimation of σ

is done overcoming the subjective influences of the investor, and it is found from a real optimization problem.

The solution σ_{opt} of such a problem is found according to the relations

$$s_{opt} \equiv \arg\left\{\min\left[0, \quad \frac{dh(s)}{ds} = 0\right]\right\} \qquad \text{or}$$

$$s_{opt} = \min\left[0, \quad \frac{d}{ds}\left(\frac{as^2 + b}{as + g}\right) = 0\right]$$

$$\frac{dh(s)}{ds} = \frac{2as(as + g) - a(as^2 + b)}{(as + g)^2} = 0.$$

It is necessary to be satisfied the condition

$$as + g = E_p > 0, \qquad a \neq 0.$$

Respectively, it holds

$$s_1^{opt} = \frac{-g + \sqrt{g^2 + ab}}{a} \quad . \tag{13}$$

This relation gives analytical way of calculation of the optimal parameter for risk preferences of the investor. For that reason the solution of the upper level optimization problem is reduced to calculation of s_{opt} according to analytical relation (13).

Example

A portfolio optimization problems is defined consisting 12 securities, traded at the Bulgarian stock exchange, $n=12$ for 12 months (June 2011 – May 2012). This problem is solved by two manners: by MATLAB – function QP and by non-iterative coordination. The matrices for the portfolio problem $Q\big|_{12x12}$ and $E\big|_{12x1}$ are defined according to the Bulgarian stock's exchange data as follows:

Q_{12}=[0.2225 0.0909 0.2582 0.2719 0.1857 0.4233 0.0251 0.0365 0.1903 0.1838 -0.5610 0.2840;
0.0909 0.0415 0.1141 0.1228 0.0876 0.1844 0.0118 0.0186 0.0786 0.0882 -0.2535 0.1267;
0.2582 0.1141 0.5003 0.3418 0.2347 0.5188 0.0312 0.0470 0.2580 0.2550 -0.3005 0.3445;
0.2719 0.1228 0.3418 0.3812 0.2725 0.5659 0.0331 0.0552 0.2255 0.2730 -0.7980 0.3865;

```
0.1857   0.0876   0.2347   0.2725   0.2041   0.4008   0.0206   0.0373   0.1445   0.2004  -0.5909   0.2731;
0.4233   0.1844   0.5188   0.5659   0.4008   0.8792   0.0479   0.0792   0.3513   0.3999  -1.2037   0.5830;
0.0251   0.0118   0.0312   0.0331   0.0206   0.0479   0.0071   0.0085   0.0260   0.0216  -0.0651   0.0358;
0.0365   0.0186   0.0470   0.0552   0.0373   0.0792   0.0085   0.0219   0.0284   0.0352  -0.1119   0.0514;
0.1903   0.0786   0.2580   0.2255   0.1445   0.3513   0.0260   0.0284   0.2297   0.1650  -0.3886   0.2563;
0.1838   0.0882   0.2550   0.2730   0.2004   0.3999   0.0216   0.0352   0.1650   0.2112  -0.5531   0.2793;
-0.5610  -0.2535  -0.3005  -0.7980  -0.5909  -1.2037  -0.0651  -0.1119  -0.3886  -0.5531   2.6543  -0.8474;
0.2840   0.1267   0.3445   0.3865   0.2731   0.5830   0.0358   0.0514   0.2563   0.2793  -0.8474   0.4177]
```

E_{12}^{T} =[3.145; 0.9318; 5.7488; 1.7963; 2.1541; 4.1983; 1.2770; 0.7481; 5.8677; 2.1795; 88.9689; 1.8094];

The efficient frontier for the optimization problem is determined for variation of sigma from 0.001 to 20 with increment of 0.1. For that case the parameter of the investor's preferences for risk S_{opt} is calculated, using relation (13), i.e. the minimal Sharp ratio risk versus return is found, according to the sequence:

1. Definition of the portfolio problem (1) with n=12;

2. Evaluation of the scalar values of the intermediate parameters $\alpha(n)$, $\beta(n)$, $\gamma(n)$ from (9);

3. Calculation of the efficient frontier with initial value s^*=0;

4. Evaluation of the portfolio parameters $V_p = V_p(\sigma^*, \alpha(n), \beta(n), \gamma(n)$, $E_p = E_p(\sigma^*, \alpha(n), \beta(n), \gamma(n))$, according to (10). In that manner, one point from the efficient frontier in the space risk/return $V_p^*(E_p^*)$ is found;

5. New value of the coefficient s is chosen, $s^{**}=s^*+0.1$. Jump to 4.

Relations "Return – s" and "Risk – s" when σ varies with 0.1 from 0.001 to 20 are given in Fig.2 and Fig.3, respectively.
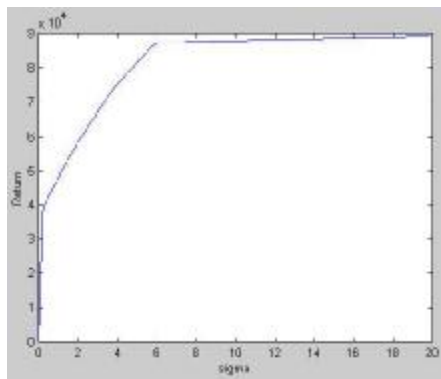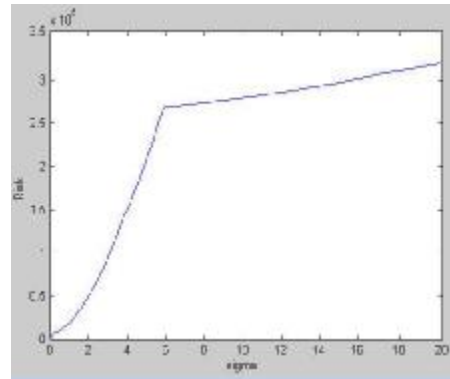
Fig. 2. Relation E(σ), σ=0.001 – 20   Fig. 3. Relation $V(\sigma)$, σ=0.001 – 20

The efficient frontier $V_p = V_p(E_p)$ in the classical case is determined point after point after numerous solutions of the portfolio optimization problem with different values of the parameter of the investor's preferences to risk σ. The efficient frontier, using MATLAB's function QP is presented in Fig.4. This curve, according to the hierarchical approach and noniterative coordination is illustrated in Fig.5. After drawing the efficient frontier, the optimal value of σ is determined. In this paper following problem (11), σ_{opt}, is calculated as solution of an upper level optimization problem according to (13). Its value is 0,000026682, given like a circle in Fig.5.

Thus, the argument s is calculated as a solution of well defined and consistent optimization problem. In comparison with the classical portfolio theory the value of s is not assessed by subjective consideration of the financial analyzer, which is an advantage of the bi-level portfolio problem.

Following figures 4 and 5, it can be seen that σ_{opt} is situated close to the origin of the graphics and it has very small value. Hence, for that case the graphical solution of the Sharp problem (14) can not be found. Thus, the analytical relation (13), gives advantages for the optimal policy of investment and quantitative evaluation of the portfolio arguments x_i – investment shares per security i and the coefficient of investor's preferences s. Hence, s is found as an argument of portfolio problem instead of coefficient, which value is chosen subjectively by the investor.
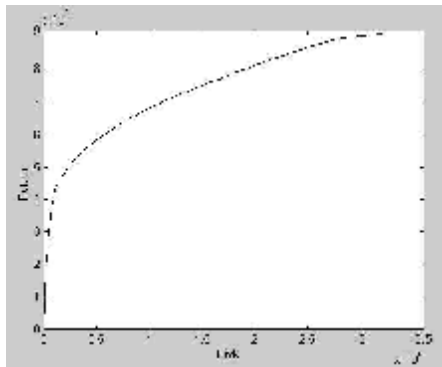
Fig. 4 Relation Return-Risk (QP function)   Fig. 5. Return-Risk (Noniterative Coordination)
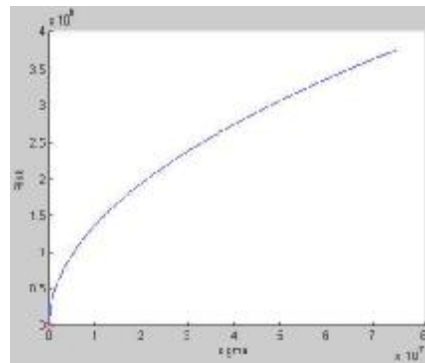
Conclusions
The classical manner of solving the portfolio optimization problem is difficult because of the unknown parameter for risk preferences s. That is why this parameter is chosen by the financial analyzer. This subjective

determination influences the portfolio optimization problem definition and solution. In the paper s is determined objectively about solution of analytically given problem, for instance the Sharp ratio, which gives the minimum relation "Risk-Return". This problem is solved at the upper level of bi-level hierarchical problem. At the lower level the classical portfolio optimization problem is solved. Previous results from so called Non-iterative coordination are applied. The main results here are in three directions: overcoming the subjective influence in definition and solution of portfolio optimization problem; calculation of s by minimizing the Sharp ratio; fast investment allocations which allow real time implementation by using the non-iterative coordination.

References
1. Bodie, Z., Kane, A., Marcus, A. *Investments*, Naturela, Sofia, 2000.
2. Campbell J., Chacko, G., Rodriguez, J., Viceira, L. Strategic Asset Allocation in a Continuous-Time VAR Model. Harvard University Cambridge, Massachusetts, 2002, 1-21.
3. Christoffersen, P. F. *Elements of financial risk management*. Vol. 1. Elsevier, 2003.
4. Kohlmann, M. Tang, S. Minimization of risk and linear quadratic optimal control theory. *SIAM J. Control optim,* Vol. 42, No. 3, 2003, 1118–1142 .
5. Mateev, M. *Analysis and assessment of investment risk*. University publisher Economy, Sofia, 2000, in Bulgarian
6. Sharpe W., Alexander, G., Bailey J. *Investments*. Prentice Hall, England Cliffs, New Jersey, 1999.
7. Sharpe, W. *Portfolio theory & Capital markets*. Mc Grow Hill, No4, 2000.
8. 8. Stoilov, T., Stoilova, K. *Noniterative Coordination in Multilevel Systems*. Kluwer Academic Publisher, Dordrecht/Boston/London, 1999.
9. Stoilova, K. *Predictive Noniterative Predictive Coordination*. Bulgarian Academy of Sciences' Publisher "M. Drinov", Sofia, 2010, 414 p., ISBN 978-954-322-268-1, in Bulgarian.

***Author*:**
Krasimira Petrova Stoilova, associated professor, DSc., PhD, Institute of Information and Communication Technologies-Bulgarian Academy of Sciences; e-mail: k.stoilova@hsi.iccs.bas.bg

EVALUATION OF ECONOMICS PARAMETERS OF BULGARIAN ASSETS BY EXCEL

Todor Stoilov

Abstract: The paper presents a numerical sequence of evaluation of characteristics of financial assets. The financial characteristics are important parameters, which are needed for the assessment and decision making of investors. These evaluations have to be performed by the financial analyzers. A prerequisite for optimal investment is the opportunity to make assessment of the securities with low cost and available software suits. The paper gives example for the evaluation of beta coefficients of securities, traded on the Bulgarian Stock Exchange. The evaluations are performed with the wide spread software suit, EXCEL.

It exist a gap between the development of models, which formalize the state and dynamics of financial markets and working calculations, algorithms and public software for assessment of the economical parameters of assets of Bulgarian Stock Exchange. The practical lack of fast calculations procedures using wide accessible tools for assessment of Bulgarian assets is a restriction for the application of formal and quantity reasonable investment solutions. For example, the usage of the theory of optimal portfolio [5] for practical applications requires execution of calculations which are not able to be always automated and implemented with wide distributed software program.

The goal of the paper is to present a sequence of calculations, performed with the widespread software program EXCEL, which can evaluate characteristics of Bulgarian assets from the Bulgarian Stock Exchange. These sequences do not require the usage of specialized and complex software products. The presented calculations in the paper can be used by the financial analyzers to make investment decisions. Practically, they can apply methods for quantitative assessment of the dynamics of the capital market.

A solution of the upper technological and practical problem is presented in the paper by defining a calculation sequence, implemented on the program suit of electronic tables – EXCEL. This solution allows fast calculation of basic and additional quantitative indices of assets, securities, financial instruments for financial analyzer or individual investor.

Analysis of Bulgarian Stock Exchange

The Bulgarian Stock Exchange appeared in the beginning of XX century. The first stock-market legislation appears in 1907 for manage-

ment of the activity of stock exchange. At the beginning the assets of 21 business entities were traded with 81 million leva basic capital. In the 30 years of last century the Bulgarian stock exchange has a significant bloom with 30 registered companies. After the Second World War the stock exchange has been interrupted till 1991 when the contemporary Trade low was accepted. "The Low for public proposition of assets" is accepted in 1999 and the stock exchange operates in correspondence with it. The current legislation provides secure environment for trade, stability at the market, keeping equality rights for all participants, transparency and irreversibility of deals, integration and globalization of the capital markets.

An important research about the Bulgarian capital market is given in [4]. The increase of the effectiveness of the capital market can be asked by improving the normative basis. However, having theoretical backgrounds about the optimal investment decisions, the participants in the stock market contribute for the increase of its importance and role. That is why the application of numerical algorithms for assessment of parameters of financial assets is important component in the application of the Capital market as an instrument for the development of the national economy.

The portfolio theory – working tool for optimization of financial investments

For the final decision of financial investment, the investor has to provide three types of activities [3]: security analysis; portfolio analysis; definition of the optimal portfolio for the investor.

Security analysis At this phase the prediction of the future characteristics of the securities, goods, and shares is required. It is necessary to be evaluated the potential income of the investment. These prognoses have to evaluate also the mutual connections between the returns of the individual assets. It is assumed that on this phase it is difficult to be defined precise quantitative relations between the parameters of the individual assets. That is why the work of the financial analyzer is compared with creative work and professional talent [5, 6].

Portfolio analysis On this stage quantitative assessment of the characteristics of the alternative portfolios is performed. It is necessary to make calculations about the average portfolio income E_p and the portfolio risk t_p, which is the standard deviation of the portfolio return. These quantitative parameters result from the forecasts the security returns, provided from the previous phase. For the portfolio analysis the financial analyzer has to make precise and adequate calculations [1].

Portfolio definition Taking into account the characteristics of the alternative portfolios about their returns E_p and risks t_p , final optimal portfolio has to be chosen, taking into account the investor's preferences for undertaking risk.

The portfolio theory is related with the second and third stage of making investment decisions. About the first stage, it is necessary to be analyzed and calculated the quantitative characteristics of the returns and risks of the individual assets and the correlations between the individual returns of the assets. For that reason it is necessary the financial analyzer to have tools and calculation algorithms which he can apply without the usage of sophisticated software.

Quantitative characteristics of the securities

The basic characteristic of the security is its return. The return is related with two quantitative parameters according to the financial theory: net present value and profit.

– *The net present value* evaluates the wealth of the securities at the current time moment. That means this is a value of 'the money now, at this moment" which the investor has. Quantitatively the Net Present Value is assessed in financial units like leva, euro.

– *The return* is a relative change of the Net Present Value. Quantitatively, it is assessed in percentages or relative units.

Nevertheless which form of return is chosen by the investor, it defines the income of the corresponding security. In that manner, the price of the security can be considered like net present value, which the investor can receive after its sell. The relative change of the security price determines the profit of the sell.

The return of each security R_i (*t*) (the real return at the moment *t*) varies during the time. To assess or forecast it influence to the portfolio, the last uses the average value of the return for the investment period E_i. This average value of the security return *i* is evaluated as weighted average sum of the real returns R_i weighted by its probability P_i :

$$E_i = \sum_{i=1}^{m} P_i R_i \qquad (1)$$

If the probabilities are equal, the relation becomes:

$$E_i = \frac{1}{m} \cdot \sum_{i=1}^{m} R_i \ , \qquad (2)$$

where *m* is the number of available values of R_i.

The average value of the security's return can be calculated by the built in EXCEL function AVERAGE(). This function executes relation (2) for the chosen set of data. Thus, the assessment of the average return of securities can be evaluated with easily available software like EXCEL.

Evaluation of the security's risk

The risk of the security is defined according to the range of values, between the real return $R_i(t)$ can take values around the average value E_i, Fig.1. This range is assessed quantitatively with the standard deviation of the dynamical function $R_i(t)$. The standard deviation t_i or dispersion t_i^2 (variation) around the average value is calculated as the average value of quadratic difference between the real value R_i and the average one E_i, weighted by the corresponding probability P_i.

$$t_i^2 = \sum_{i=1}^{m} P_i (R_i - E_i)^2 \ . \tag{3}$$

For the case of equal probability of a set of data the relation becomes

$$t_i^2 = \frac{1}{m} \cdot \sum_{i=1}^{m} (R_i - E_i)^2 \ , \tag{4}$$

where m is the number of the values from the data set.
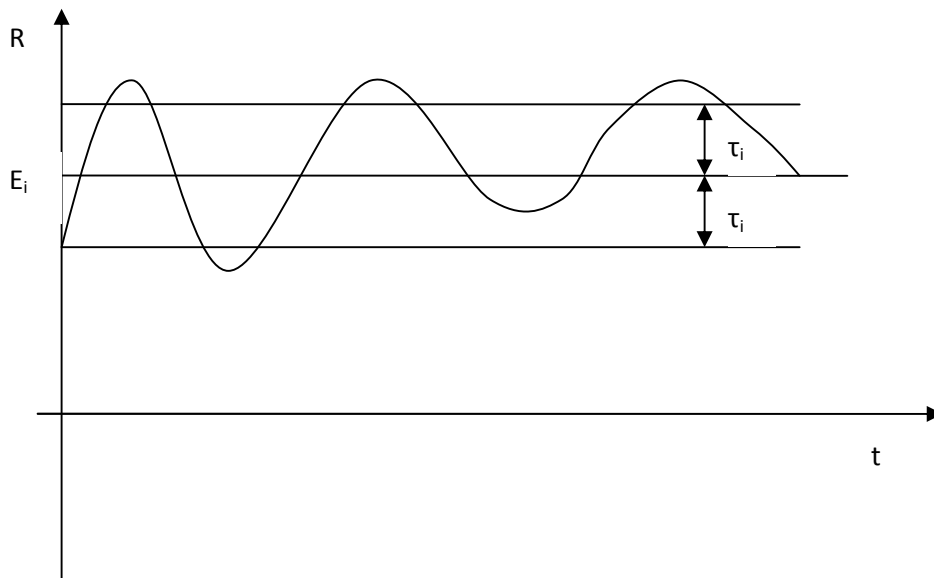
Fig.1. Description of the average value of E_i and risk t_I of financial security

Relations between the securities returns

Important peculiarity of the portfolio theory is that it has to take into the mutual influences between the returns of securities. These relations are determined quantitatively by correlation coefficients or co-variations [2]. These coefficients provide assessment about the range of influences of the return of one security to the return of other security [6]. The correlation coefficient is calculated as weighted mean value of relation:

$$r_{jk} = \sum P_r(R_j, R_k) \frac{(R_j - E_j)}{t_j} \times \frac{(R_k - E_k)}{t_k} , \qquad (5)$$

where $P_r(R_j, R_k)$ is the probability for simultaneous existence of the couple R_j, R_k.

Instead of the correlation coefficients, for the portfolio optimization are used the co-variations between the securities returns, which are calculated more easy as

$$c_{ij} = \sum_{l=1}^{\min(m_i.m_j)} \sum_{p=1}^{\min(m_i.m_j)} P_r(R_l, R_p)[(R_l - E_i)(R_p - E_j)], \qquad (6)$$

where m_i, m_j is the number of the available values of R_i and R_j.

The relation between the co-variation and correlation is

$$c_{jk} = r_{jk} t_j t_k \quad , \qquad (7)$$

where R_i is the real value of the return of the security i; E_i is the average value of return; cov(.) is the co-variation matrix containing c_{ij}.

The calculation of the set of co-variation coefficients can be performed with the built in EXCEL function COVAR(). These coefficients define a matrix with symmetric structure. This peculiarity simplifies the number of calculations, which have to be done only for the components of the matrix above the main diagonal. The diagonal coefficients of the co-variation matrix give the dispersion of each security, t_i^2, i=1,…,N, N – number of securities. In that manner for the calculation of the security's risk it is not necessary to be used another specialized function from EXCEL.

Example for calculation of beta-coefficients of securities

An example for the application of the above mention built-in function AVERAGE() and COVAR() is given here, by means to evaluate the important financial parameters, named beta coefficients b_i.

The theory of capital market is an extension of the portfolio theory [5]. This theory develops formal relations between the characteristics of

the individual securities and the market as a whole. The integral assessment of the markets is performed with introduction of indices like Dow Jones, Nasdaq. Currently, the Bulgarian Stock exchange also calculates indices like SOFIX, BG40 etc. The capital market theory applies the values of the indices for the assessment of the market behavior and the individual characteristics of the assets. An important relation, which is derived, is the so called characteristic line (*HL*) of security *i*, $HL = R_i(R_M)$. This line gives linear relation between the individual security return and the value of the index. Respectively, this line has to consider that if the current market index R_M takes value equal to its average one E_M, then the characteristic line has to give real value of the security return R_i, equal to its average value E_i, i.e. if $R_M = E_M$, then $R_i = E_i$. When $R_M = 0$, then $E_i = p$, where *p* is the average market riskless loan. The slope of the line *HL* defines the correlation of the security risk with the market behavior and it is evaluated as

$$C_{iM} = \sum P_r(R_M)(R_i - E_i)(R_M - E_M),$$

where $P_r(R_M)$ is the probability the market index to have a value R_M. The characteristic line *HL* is derived by linear relation between R_i and R_M or

$$R_i = a_i + b_i R_M,$$

where a_i and b_i are unknown parameters which have to be determined.

For the value E_M the line HL must give the value E_i, or

$$E_i = a_i + b_i E_M.$$

After a substitution of E_i and R_i in C$_{iM}$ it follows:

$$C_{iM} = \sum P_r(R_M)(R_i - E_i)(R_M - E_M) = \sum P_r(R_M)\left[\underbrace{(a_i + b_i R_M)}_{R_i} - \underbrace{(a_i + b_i E_M)}_{E_i} \right](R_M - E_M) =$$

$$= b_i \sum P_r(R_M)(R_M - E_M)(R_M - E_M) = b_i \underbrace{\sum P_r(R_M)(R_M - E_M)^2}_{t_M^2} = b_i T_M^2$$

where

$$\sum P_r(R_M)(R_M - E_M)^2 = t_M^2.$$

From this relation it follows that the coefficient b_i can be calculated from C_{iM}, and t_M

$$b_i = \frac{C_{iM}}{t_M^2}.$$

88

The coefficient b_i defines the slope of the characteristic line HL of the ith security. It depends on correlation parameter C_{iM} between the ith security and the market variation. Security with a value $b_i < 1$ is called "defensive". Increase with 1% of the market index R_M results to increase of R_i less than 1%. In the opposite case, 1% decrease of the market index R_M will result in less than 1 % decrease of the income R_i of the security i. In that manner the investor is "protected" against the deterioration of the market. A security with coefficient $b_i > 1$ is called "aggressive". The decrease with 1% of the market index R_M decreases more the return R_i of security i. In the opposite case, the increase of the market R_M leads to stronger increase of the income R_i of security i. The identification and evaluation of the coefficients b_i is a widespread practice for assessment of the influence of the security i returns according to the market behavior.

The paper gives an example how to evaluate these important parameters with the application of EXCEL suit. For the example it has been used real data taken from the site of the Bulgarian Stock Exchange (www.bse-sofis.bg), both with the values of the index SOFIX for the first 5 months of 2012. Five 5 securities have been chosen from the official segment of the market, Table1.

Table 1 Prices of securities from the Bulgarian Stock Exchange

Exchange code	Emitent	Date/price	Date/price	Date/price	Date/price	Date/price
		3.1.2012	3.2.2012	2.3.2012	3.4.2012	3.5.2012
3JR	Sopharma AD	3.154	2.957	2.948	2.799	2.571
4CF	Central Cooperative Bank	0.831	0.848	0.786	0.765	0.758
5F4	First Investment Bank	1.834	1.979	1.95	1.83	1.81
5MB	Monbat	5.366	5.325	5.279	5.211	5.11
6C4	Chimimport	1.389	1.431	1.418	1.333	1.259
SOFIX	318.03	318.03	317.73	313.38	309.37	303.32

These data are written on EXCEL sheet in sells B3:B7 to G3:G7, fig.2. The average return of the securities is calculated at row 10. In cell

B10 is written the formula =AVERAGE(B3:B7). Respectively, in the last cell G10 is written =AVERAGE(G3:G7). In that manner the average returns E_i have been evaluated.

In cells B14:B19 to G14:G19 are calculated the co-variation coefficients of the security returns according to (6). The formula =COVAR(B3:B7,C3:C7) is used, which defines the values in cells C14 and B15, tacking into account that the co-variation matrix is a symmetric one. On the main diagonal are the variations of the security risks, according to (4). Thus, the cell C14 contains t_i^2. For the value of t_i in a cell B21 is written the formula =SQRT(B14).

	A	B	C	D	E	F	G
1	Банков код	3JR	4CF	5F4	5MB	6C4	SOFIX
2		BGL	BGL	BGL	BGL	BGL	
3	3.1.2012	3,154	0,831	1,834	5,366	1,389	318,030
4	3.2.2012	2,957	0,848	1,979	5,325	1,431	317,730
5	2.3.2012	2,848	0,786	1,950	5,275	1,418	313,380
6	3.4.2012	2,799	0,765	1,830	5,211	1,333	309,370
7	3.5.2012	2,571	0,758	1,810	5,110	1,253	303,320
8							
9	Средна доходност	E1	E2	E3	E4	E5	ES
10		2,866	0,798	1,8806	5,2574	1,3648	312,366
11							
12	Ковариационна матрица						
13		E1	E2	E3	E4	E5	ES
14	E1	0,03661	0,005656	0,0037	0,016741	0,00943	30,53754
15	E2	0,00566	0,001284	0,001409	0,002836	0,001798	0,18073
16	E3	0,0037	0,001409	0,004843	0,003113	0,003716	0,224898
17	E4	0,01674	0,002836	0,003113	0,008111	0,005199	0,493592
18	E5	0,00943	0,001798	0,003716	0,005199	0,004262	0,330545
19	ES	30,5375	0,18073	0,224898	0,493592	0,330545	30,5375
20							
21	Риск на актива	0,19134	0,035836	0,069592	0,090059	0,065282	5,526079
22							
23	beta=G14/B14	834,073	140,7295	46,43744	60,85729	77,56073	
24							

Fig.2. Calculation of the coefficient beta by EXCEL

For calculation of the coefficients b_i it has been used the values of the co-variation coefficients from column G14 to G18. In a cell G14 is evaluated the co-variation coefficient $c_{1,Sofix}$; in a cell G15 is $c_{2,Sofix}$, etc.

Finally, the coefficient b_1 is evaluated as

$$b_1 = c_{1,Sofix} / t_{12}.$$

The needed values are taken respectively from cell G14 and from the main diagonal B14 or b_1=G14/B14. By the same way, the coefficient b_2 is

$$b_2 = c_{2,Sofix} / t_{22},$$

where the corresponding values are taken respectively from cells G15 and from the main diagonal or b_2=G15/B15, etc.

The example illustrates the manual usage of electronic tables EXCEL without a need for the application of special and complex programs for evaluation of financial characteristics of the securities.

Conclusion

The assessment of securities characteristics, traded at the stock exchange, requires their calculations to be done in real time. The usage of specialized software can perform fast and quantitative financial analysis. However, the limited distribution of such products is obstacle for the wider application of quantitative methods in investment decision making. This paper presents an example how to evaluate important quantitative characteristics of securities, using wide available software suit as EXCEL.

References
1. Body, Kane, Marcus. *Investments*, Naturela, Sofia, 2000, p.906.
2. Campbell, J., G. Chacko, J. Rodriguez, L. Viceira. Starategic asset allocation in a continuous-time VAR model. – In: Harvard Institute of Economic Research, Harvard University Cambridge, Massachusetts, September 2002, 1-21
3. Mateev, M. *Analysis and appreciation of the risk choosing investments.* University Publisher "Economics", Sofia, 2000, 223 p. (in Bulgarian).
4. Moravenov, M. *Development of Fond market in Bulgaria and directions for its effectiveness improvement*, Ciella, 2004, 386 p. (in Bulgarian)
5. Sharpe W. *Portfolio theory and capital markets*, Mc Graw-Hill, New York, London, Tokyo, 2000, p.316.
6. Sharpe W., Alexander G., Bailey J. *Investments.* Prentice Hall, England Cliffs, New Jersey, 1999, 5 edition, p.1058.

Author: professor DSc. **Todor Stoilov**, Institute of Information and Communication Technologies-Bulgarian Academy od Sciences, 1113, Acad. G. Bonchev str., bl.2, tel: +359 2 979 27 74;
e-mail: todor@hsi.iccs.bas.bg

WEB ANALYTIC TOOLS FOR MEASURING SEO EFFICIENCY

Antonina Ivanova

Abstract: For successful performance in Internet the company's websites are optimized for search engines. There are various tools that facilitate this process. How effective SEO is can be defined with different indicators and tools. The article discusses some tools that illustrate the results of optimization and provide opportunities for analysis and assessment.

Key words: SEO, e-business, Webalizer, AWstats, Google Analytics, Google AdWords

Introduction

Nowadays companies' performance in Internet is essential to business success. In order for potential customers to reach the websites they should be detectable in searches. Analyses show that the 93% of users view only the first two pages of search results [6]. Most people prefer to use another search or another search engine rather than continue to consider sites beyond the first 10 results. It is therefore important to optimize sites for search engines (Search Engine Optimization).

Search Engine Optimization (SEO) is a plan implemented to increase a site's visibility. It involves selecting keywords that will generate targeted traffic, researching search engine policies and guidelines, and understanding customers. In addition, SEO involves dealing with design and content problems that might hinder a site's results ranking. SEO is the practice of optimizing web pages in a way that improves their ranking in the web search results, which are also known as natural or organic results because they are supposed to reflect relevancy in searchers' standard [1].

For a good SEO there are different approaches but nowadays content creation is the most important step for good site positioning but most difficult [2]. In this Search Marketing Benchmark Report is shown that when done well, keyword and keyphrase research delivers results and creates a competitive advantage, explaining its popular levels of usage. Creating title tags is another keyword-related and commonly used tactic – one which marketers considered to be equally effective. Good content creates visits and attracts links. For this reason, marketers who commit to the effort required in creating quality content can improve their SEO positions. Another tactics included in the survey are external links, url structure, meta description tags, blogging, social integration, internal linking, competitor benchmarking. SEO is not an one-time job. It requires

trial and error testing, regular updates, and periodic performance enhancements so the site will continue to appear high in the results. Because of the demands of SEO, many companies outsource these needs to individuals or businesses that specialize in providing those types of services.

How effective SEO is can be defined with different indicators and tools. The paper discusses 3 web analytic tools that illustrate the results of optimization and provide opportunities for analysis and assessment.

Background definitions for SEO measuring parameters and web analytics

Some of the most commonly-used website traffic analysis and SEO measuring terms are listed below [11], [11]:

Hit. A hit is a single request made to a web server for an object on the web site. The object can be an HTML file (i.e. a web page), a graphic image, or any other embedded object in your web pages, such as a sound file.

Visit. A visit to a web site represents one unique viewer (or search engine) who has visited the site. For example, one person visits the site and views one of the web pages. The web page has 10 unique graphic images on it (a masthead, bullet points, some photos, etc.) and a sound file. The total number of hits would be as follows:

Number of Unique Visitors	1
Number of HTML documents downloaded	1
Number of unique graphic images downloaded	10
Number of embedded files downloaded	1
Total Number of Hits	**13**

1 visitor viewing 1 web page = 13 hits

Page. A page is a successful HTTP request for a resource that constitutes primary website's content. Pages are usually identified by a file extension (e.g. .html, .php, .asp, etc.) or by a missing extension, in which case the subject of the HTTP request is considered a directory and the default page for this directory is served.

File. Each successful HTTP request is counted as a file.

Visitor. A visitor is the actual person browsing the website. A typical website serves content to anonymous visitors and cannot associate visitors with the actual person browsing the website. Visitor identification may be based on their IP address or an HTTP cookie. The former approach is simple to implement, but results in all visitors browsing the same website from behind a firewall counted as a single visitor. The latter approach requires special configuration of the web server (i.e. to log HTTP cookies) and is more expensive to implement.

Visit. A visit is a series of HTTP requests submitted by a visitor with the maximum time between requests not exceeding a certain amount configured by the webmaster, which is typically set at 30 minutes. For example, if a visitor requested page A, then in 10 minutes page B and then in 40 minutes page C, then this visitor has generated two visits, one when pages A and B were requested and another when the page C was requested.

Host. In general, a host is the visitor's machine running the browser. Hosts are often identified by IP addresses or domain *names*. Those web traffic analysis tools that use IP addresses to identify visitors use the words hosts, domain names and IP addresses interchangeably.

User Agent. User agent is a synonym for a web browser.

Backlinks. Backlinks are incoming links to a website or web page.

Bounce Rate. Recorded as a user reaches a site and leaves without navigating

Conversion rate is the ratio of visitors who convert casual content views or website visits into desired actions based on subtle or direct requests from marketers, advertisers, and content creators.

Tools for tracking SEO efficiency and website analytics
Webaliser

The Webalizer is a GPL application that generates web pages of analysis, from access and usage logs and is a commonly used web server administration tools. It was initiated by Bradford L. Barrett in 1997. Statistics commonly reported by Webalizer include: hits; visits; referrers; the visitors' countries; and the amount of data downloaded. These statistics can be viewed graphically and presented by different time frames, such as per day, hour, or month.

Website traffic analysis is produced by grouping and aggregating various data items captured by the web server in the form of log files while the website visitor is browsing the website.

The Webalizer analyzes web server log files, extracting such items as client's IP addresses, URL paths, processing times, user agents, referrers, etc. and grouping them in order to produce HTML reports. The Webalizer may be configured through parameters of a configuration file. There are over one hundred available configuration parameters, which make The Webalizer a highly-configurable web traffic analysis application.
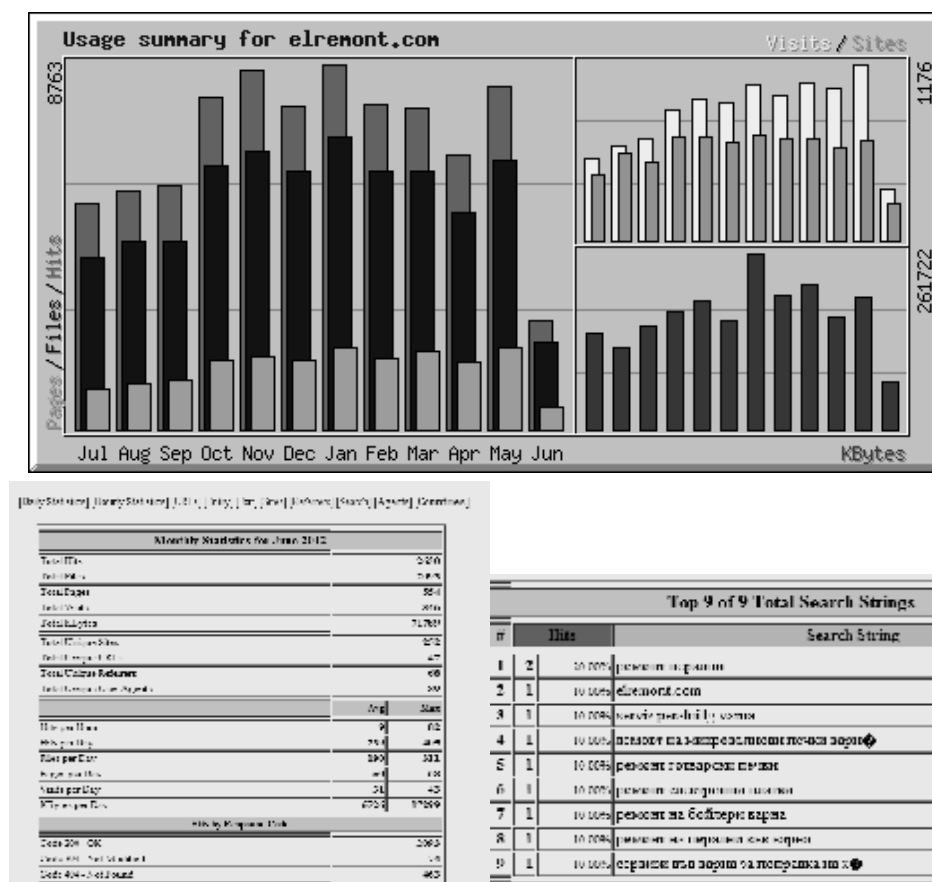
Fig. 1. Webalizer reports

Reports

By default, The Webalizer produces two kinds of reports - a yearly summary report and a detailed monthly report, one for each analyzed month. The yearly summary report provides such information as the number of hits, file and page requests, hosts and visits, as well as daily averages of these counters for each month. The report is accompanied by a yearly summary graph.

Each of the monthly reports is generated as a single HTML page containing a monthly summary report (listing the overall number of hits, file and page requests, visits, hosts, etc.), a daily report (grouping these counters for each of the days of the month), an aggregated hourly report (grouping counters for the same hour of each day together), a URL report

(grouping collected information by URL), a host report (by IP address), website entry and exit URL reports (showing most common first and last visit URLs), a referrer report (grouping the referring third-party URLs leading to the analyzed website), a search string report (grouping items by search terms used in such search engines as Google), a user agent report (grouping by the browser type) and a country report (grouping by the host's country of origin).

Each of the standard HTML reports described above lists only top entries for each item (e.g. top 20 URLs). The actual number of lines for each of the reports is controlled by configuration. The Webalizer may also be configured to produce a separate report for each of the items, which will list every single item, such as all website visitors, all requested URLs, etc.

Although easy of use Webalizer has some *weaknesses.* Generated statistics do not differentiate between human visitors and robots. As a result all reported metrics are higher than those due to people alone. Many webmasters claim that webalizer produces highly unrealistic figures of visits, which are sometimes 200 to 900% higher than the data produced by javascript based web statistics such as Google Analytics. No query string analysis. Dynamic generated websites can not be listed separately (e.g. PHP pages with arguments).

AWStats

AWStats is short for Advanced Web Statistics and is free software distributed under the GNU General Public License [4]. AWStats is powerful log analyzer which creates advanced web, ftp, mail and streaming server statistics reports based on the rich data contained in server logs. Data is graphically presented in easy to read web pages which can represent the results of SEO process. AWStats development started in 1997 and is still developed today by same author (Laurent Destailleur).

Through the use of intermediary data base files, AWStats is able to quickly process large log files, as often desired. With support for both standard and custom log format definitions, AWStats can analyze log files from Apache (NCSA combined/XLF/ELF or common/CLF log format), Microsoft's IIS (W3C log format), WebStar and most web, proxy, wap and streaming media servers as well as ftp and mail server logs.

Reports

AWStats' reports include a wide range of information on the web site usage.

• Number of visits, and number of unique visitors, visit duration and latest visits.

• Authenticated Users, and latest authenticated visits.

• Usage by Months, Days of week and Hours of the day (pages, hits, KB).

• Domains/countries (and regions, cities and ISP with Maxmind proprietary geo databases) of visitor's hosts (pages, hits, KB, 269 domains/countries detected), hosts list, latest visits and unresolved IP addresses list.

• Most viewed, Entry and Exit pages, most commonly requested File types.

• Visitor's Browsers (pages, hits, KB for each browser, each version, 123 browsers detected: Web, Wap, Streaming Media browsers..., around 482 with the "phone browsers" database).

• Visitor's Operating Systems (pages, hits, KB for each OS, 45 OS detected).

• Robots visits, including search engine crawlers (381 robots detected).

• Search engines, Keywords and Phrases used to find the site (The 122 most famous search engines are detected like Yahoo, Google, Altavista, etc...)

• HTTP Errors (Page Not Found with latest referrer, ...).

• User defined reports based on url, referrer (referer) fields extend AWStats' capabilities to provide even greater technical and marketing information.

• Number of times the site is added to Bookmarks / Favorites.

• Ratio of integrated Browser Support for: Java, Flash, Real G2 player, Quicktime reader, PDF reader, WMA reader (as above, requires insertion of HTML tags in site's home page).
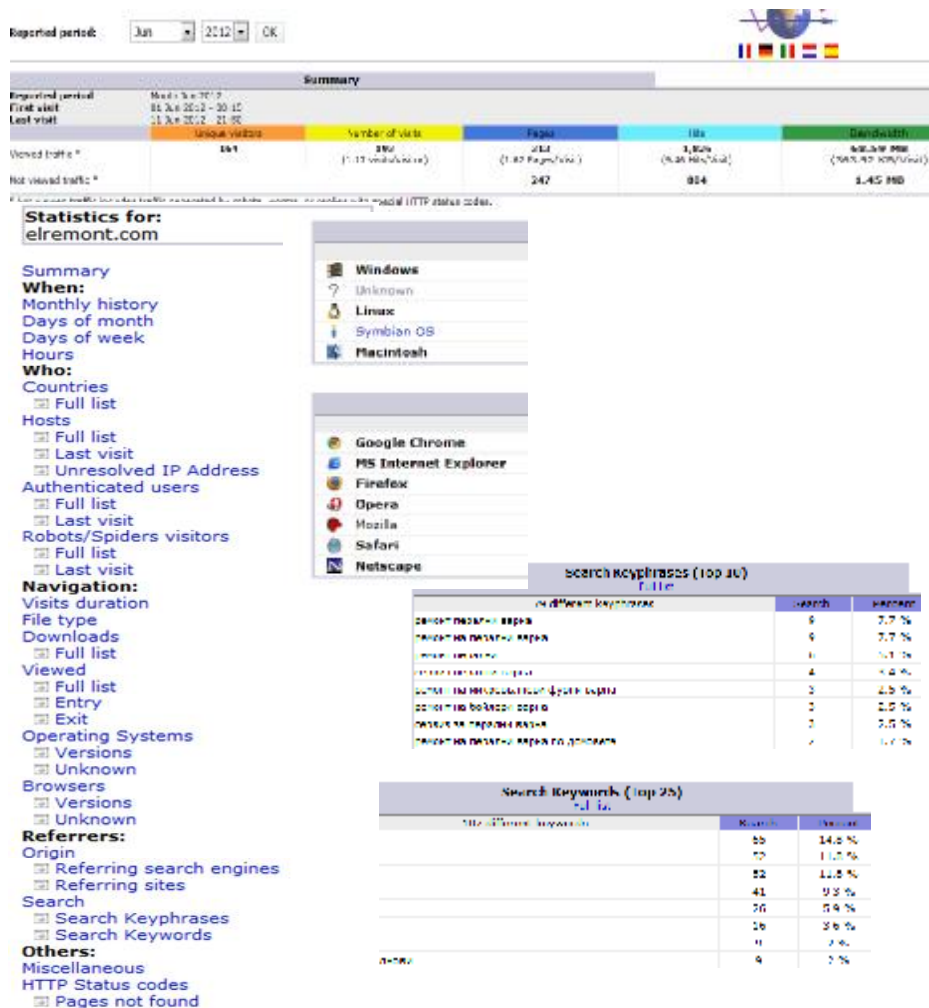
Fig. 2. AWstats report

Google Analytics

Google offers a set of powerful tools to track results of SEO in websites. The most popular of them is Google Analytics. It is a free service offered by Google that generates detailed statistics about the visitors to a website. It is the most widely used website statistics service.

GA can track visitors from all referrers, including search engines, display advertising, pay-per-click networks, e-mail marketing and digital collateral such as links within PDF documents.

GA's approach is to show high-level, dashboard-type data for the casual user, and more in-depth data further into the report set. Through

the use of GA analysis, poor performing pages can be identified using techniques such as funnel visualization, where visitors came from (referrers), how long they stayed and their geographical position. It also provides more advanced features, including custom visitor segmentation.

Users can officially add up to 50 site profiles. Each profile generally corresponds to one website. It is limited to sites which have a traffic of fewer than 5 million pageviews per month (roughly 2 pageviews per second), unless the site is linked to an AdWords campaign.[5]

Google Analytics is implemented by including what is known as a "page tag". This is referred to as the Google Analytics Tracking Code (GATC) and is a snippet of JavaScript code that the user adds onto every page of his or her website. This code collects visitor data and sends it to a Google data collection server .

Reports

The dashboard gives a basic summary of the results, including daily visits, average time on site, and traffic types. The user can choose standard or custom reporting.

Audience Overview

The Audience section includes the behaviors, demographics and technologies of your visitors. The analyst can click on the links in the side navigation to drill down to more detailed information for each area. Some important areas to analyze on the Audience Dashboard include:

• How many pages are viewed each visit. Knowing how many pages are viewed at each visit in combination with other metrics can be a good analysis tool. This does not apply in some cases, including Blogs, as visitors commonly visit only one page.

• How long visitors stayed on the site.

• New vs. Returning Visitors. This is a basic metric that gives insight into whether getting unique visitors or repeat visitors. The Behavior > New vs. Returning Report lefthand side allows to go more in depth. This can be used to determine how many new visitors are coming to the site. On the other hand it also helps determine how many visitors are returning. This metric should be compared with other metrics to determine a course of action.

Advertising Overview

The Advertising section gives a closer look at the Google AdWords program which is integrated with Google Analytics. Users can review online campaigns by tracking landing page quality and conversions (goals). Goals might include sales, lead generation, viewing a specific page, or downloading a particular file. These can also be monetized. By using GoogleAnalytics, marketers can determine which ads are

performing, and which are not, providing the information to optimise or cull campaigns. Analyzing these reports allows to determine the effectiveness of the keywords, landing pages & campaigns.

Traffic Sources Overview

The Traffic Sources section indicates how visitors found the site. In this section it is possibleto evaluate & compare the effectiveness of the search, referral, direct & campaign traffic. A more detailed report can include average time on site, bounce rate, & pages per visit. Compare conversion & Ecommerce effectiveness of each traffic source by analyzing goal & Ecommerce conversion rates, per visit goal values & revenue.

The Traffic Sources reports can be used to helps making decisions about where to focus advertising and site-optimization efforts.   The following metrics can be used as good determinates for qualified traffic:

- Bounce rate is the percentage of visits to your site where the visitor leaves after viewing only one page. Bounce rate is a good indicator of qualified traffic except in a few cases like if the site is a Blog. In cases like this, it is very common for visitors to only view one page and leave. Regardless of how long they look at the page, the bounce rate would still be high.

- Time on site can also be a good indicator of what traffic is qualified. If a visitor is not spending a lot of time on the site, it may indicate that the visitor is not seeing what they are looking for. Based on Time on Site numbers, you may want to consider possible plans of action

Enabling Google Webmaster Tools within Google Analytics

Google Webmaster Tools is a free product that provides with detailed reports about site's visibility on Google. If Webmaster Tools is set up for the site, it is possible to get Search Engine Optimization data from Webmaster Tools in the Analytics reports.

Content Overview

The Content section is designed to help toevaluate & improve the content on theWebsite. The overview displays the number of pages that were viewed on the site, along with a list of the most visit pages. The Site Content reports can be used to indentify exit pages & pages with high bounce rates or low average time on site. Analyzing these reports can help to identify pages that may need to be rewritten or reorganized.

Conversions Overview

Conversions represent the completion of an activity on the site that is important to the business.

Goal Conversions

Within the Goals reports provide the number of completed goals & the conversion rate for each of the goals which was set-up. Goals can be

combined with other metrics to aid in determining what is prompting the goal completion.

Ecommerce Conversions

Ecommerce tracking is Javascript code on the shopping cart or site that collects information about each transaction, including which product(s) were purchased, purchase amounts, billing city, state, and country. Tracking Ecommerce transactions also allows to see the value of traffic by geography, keyword, referral, campaign, and many other dimensions.
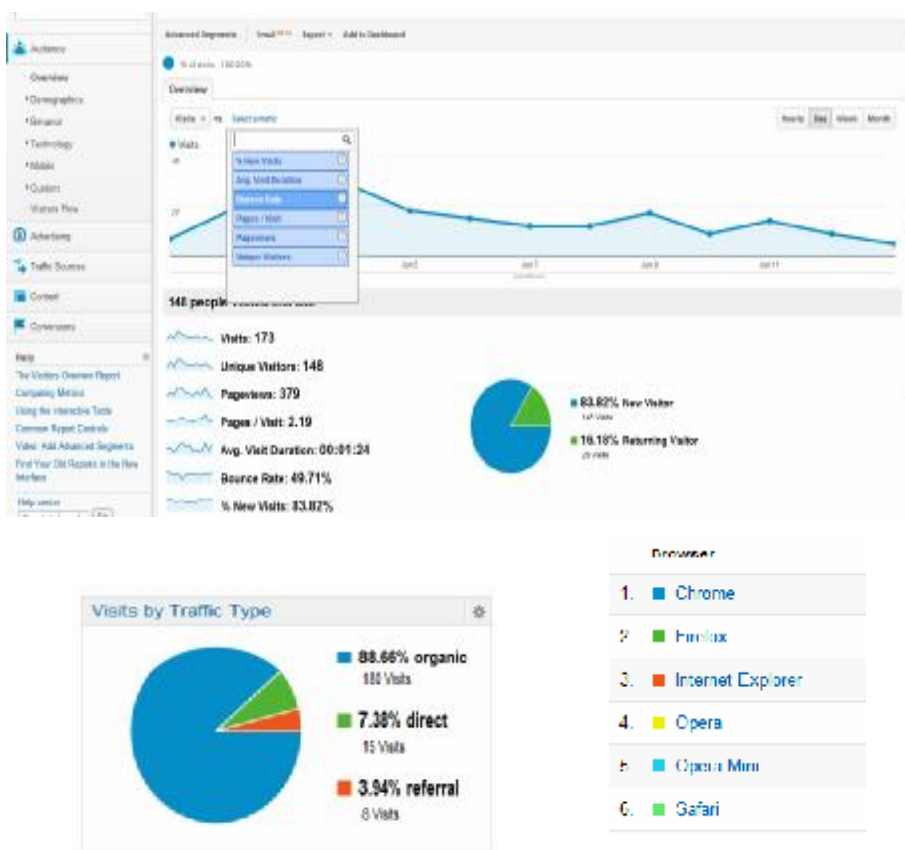
Fig. 3. Google Analytics reports

Conclusions

Web masters and analyzers have a wide range of tools for evaluating SEO process in web sites. In this paper only 3 of them were presented. The observation of one site – http://www.elremont.com for the

same period with the three tools provided different report values. Webalizer can show results with up to 900% higher values for visits and it is recommended to use t is tool combined with another one. The results of AWstats and Google Analytics are not equal too but closer each to one. AWstats counts more visits and visitors than Google Analytics. Both tools show different results for browsers statistics. Google Analytics provides more complex report opportunities and is integrated with other Google instruments. This makes it preferable for website analytics and measuring SEO results.

References
1. Bo Xing, Zhangxi Lin. The Impact of Search Engine Optimization on Online Advertising Market. Conference'04, Month 1–2, 2004, Copyright 2004, ACM 1-58113-000-0/00/0004, page 519-529, May 7, 2011.
2. Bower, K. 2012 Search Marketing Benchmark Report – SEO Edition MarketingSherpa LLC, a MECLABS Group Company, 2011.
3. Clifton, Brian. Advanced Web Metrics with Google Analytics, Third Edition. John Wiley & Sons, 2012.
4. http://awstats.sourceforge.net/docs/awstats_what.html
5. http://seo.ezinemark.com/how-to-assess-the-effectivity-of-your-seo-company-7d3309cff62c.html
6. http://seom.bg/
7. http://www.webalizer.org
8. https://freekeywords.wordtracker.com
9. https://www.google.com/analytics
10. Thakur A., A. L. Sangal, H. Bindra Quantitative Measurement and Comparison of Effects of Various Search Engine Optimization Parameters on Alexa Traffic Rank. In: International Journal of Computer Applications (0975 – 8887), Volume 26– No.5, July 2011.
11. Web site promotion scam: using hits and visits in online. Marketing. http://www.grantasticdesigns.com/hitsvisits.html
12. Иванов, Андрей. Оценка эффективности продвижения сайта. Выпуск 27. Классическая задача оптимизации. Метод учета эффективности SEO, Optimization.ru, 2005

Author:
Antonina Ivanova Ivanova, Assist. Prof.
Varna Free University, Chaika Resort Varna 9007, Bulgaria
+ 359 052 359574
krumova@vfu.bg

INNOVATIVE PRINCIPLES IN BUILDING THE I.DB.I. ARTERY (INTERNATIONAL DATABASE FOR EAST CHRISTIAN ICON ART)

Radoslav Yoshinov, Ivan Garnizov

Laboratory of Telematics to the Bulgarian Academy of Sciences

Abstract. The East Christian Iconography, also described as 'Orthodox Christian Icon art' comprises not only the well-known icons on panel, but also the monumental art in situ like wall paintings and mosaics, as well as the small copper icons which the orthodox believer keeps in his pocket for veneration and protection. It has been a neglected area in the digitized documentation and registration of plastic arts. At present multimedia information (text, high quality images, video, 3D models etc.) of the different forms of icon-art are registered and classified in 'International DataBase on East Christian Icon art' (I.DB.I.) Artery. The aim of the Artery is through compilation of a comprehensive basic system for documentation and registration of the East-Christian cultural heritage in possession of museums, churches, institutes and private persons guiding the whole life cycle from information store to extract, retrieval and publishing in any kind, to increase accessibility for the European citizen of this human cultural heritage. Currently the I.DB.I. Artery is based on the existing "Innoclature" description scheme, which has been developed in the seventies of the 20-th century carrying the technical specifications of its time. Migrating the I.DB.I. Artery from the existing "Innoclature" description scheme to the innovative Metadata Encoding and Transmission Schema as a standard for encoding descriptive, administrative, and structural metadata regarding resources within a digital library will bring the I.DB.I. on federated digital content level.

Key words: East Christian icon art, digital libraries, database, cognitive system

INTRODUCTION

The East Christian Iconography, also described as 'Orthodox Christian Icon art' comprises not only the well-known icons on panel, but also the monumental art in situ like wall paintings and mosaics, as well as the small copper icons which the orthodox believer keeps in his pocket for veneration and protection. Techniques, materials and iconography of the icons in all forms, are in principle the same. The invariability of presentations of the Holy Image (God, saints, and events) enables the immediate recognition of any icon, from any period, from any region in the long and widespread culture of East Christian Orthodoxy. Some icons, as images of the Holy, were pretended to be made 'by human hands driven by God' in the orthodox iconography and might be compared with the Gospels, said to be the Word of God, given by God: untouchable, invariable. But as they were translated in many languages which evolve in time over the hundreds of years, the canonical images

also obtain "regional" style changing in the course of time, assuming the individuality of their creators opposed to the canonical rules which makes their icon a masterpiece of art. Still the basic iconography, the standard types stayed the same. Taking all this into account during the UNESCO/ODIGIA conference for establishing an International DataBase on East Christian Icon art (I.DB.I.) in 1999, an international group of museum directors and experts decided to start the construction of a highly visual and interactive documentation and registration system for representations of Eastern Christian art in all media, in which highly scattered cultural information will be brought together using a common information system providing tools for navigation through this 'virtual encyclopedia' in all conceivable ways. This system, called I.DB.I. Artery is a compilation of a comprehensive basic system for documentation and registration of the East-Christian cultural heritage in possession of museums, churches, institutes and private persons, taking into account Intellectual Property Right (IPR) associated with the processed objects. It should guide the whole life cycle from information store to extract, retrieval and publishing in any kind. Any object, entering the system has to be put automatically in the right place between the others. For this analysis of: I. Which standard type of icon (plastic art) it is; II. Which (standard) variation on that type; III. Which individual specifics it shows; IV. Where might it come from; V. What could be the period is needed. Main measure points are landscape, architecture, figure(s), dress, and physiognomy. For I and II definition of icon pattern is needed. For III, IV and V detail-pattern analysis of the object is needed. The development of the I.DB.I. Artery became the base for registration, documentation and educational research of innumerable important works of art. Here on Fig. 1 is shown the general idea of the I.DB.I. Artery.
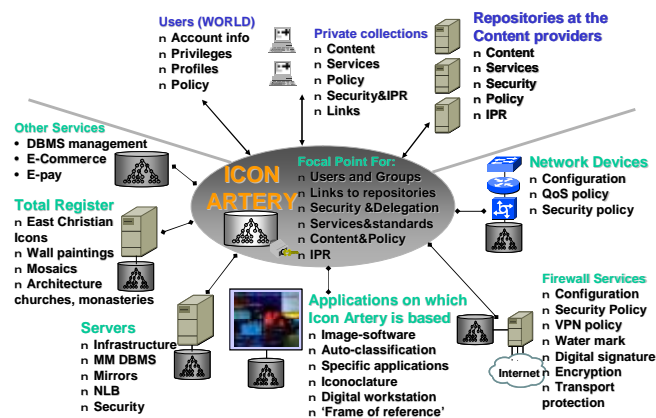
Fig. 1. Hierarchical solution of the portfolio optimization problem

An image-analyzing tool is extended in 3D in addition to the existing basic search and retrieval tools to provide optimal accessibility to the content of the system. This image-analyzing application will have the analytic facilities to discern color, shapes 2 and 3D, different techniques and recurring patterns. This will enable to fit in works of art in an art historic context, discovering unexpected links to existing and still hidden and often even unknown treasures of Eastern Christian art. The compilation of a reference-register in a user-friendly information system containing specific, distinctive icons and plastic art objects is essential to make a reliable analysis using methodologies and tools for effectively and efficiently managing large, federated, digital document collections obeying to the emerging standards for document management (XML, RDF, DOM, XQL or XML-QL etc.), and providing at the same time the sound object management capabilities. The emphasis is also on the effective integration, which will take advantage of the more detailed knowledge about the document contents due to the XML description (XML Schema, DTD) for metadata extraction, in order to achieve better categorization and content addressability. The selected meta-data model is of particular importance as it influences the choice of attributes to describe an object. So here the important technology shift is passing from the existing Innoclature description scheme to the innovative Metadata Encoding and Transmission Schema (METS)[8] as a standard for encoding descriptive, administrative, and structural metadata regarding resources within a digital library. This will bring the I.DB.I. on federated digital content level. These innovative technological principles for extending the I.DB.I. Artery were partially included in the development of BG-Sacra project in Bulgaria [9] [11]

Innovation principles to the I.DB.I.

Technological innovation principles
Now the register of I.DB.I., based on the Innoclature and using the relational DBMSs are basic system for documentation and registration of the East-Christian cultural heritage in possession of museums, churches, institutes and private persons. The important question here is to embed and migrate all the existing content and functionalities of this system into a Digital Library (DL)[1] with tight integration to the standards of the document community (XML, RDF, DOM, XQL) which will allow document interfacing to the outside world with enhanced functionality, offering document component sharing, versioning, security, concurrency etc. Important capabilities of the so created I.DB.I. Digital Library will also be the interconnection with other libraries and the federation through trustable connections (XLink) and capabilities to describe the contents of the des-

tination documents as well as the charging involved for the access. For this wrappers for conversion and methodologies for uniform content addressability must be provided. Methodologies and tools are developed for automatic capturing and adaptation of metadata from the Innoclatura application, with flexible data placement according to the document DTDs, types, size, etc. The place from where the metadata is extracted, the importance of the various parts of the documents and the rules and statistics followed will depend on the particular DTD and document type. Documents will be automatically categorized into dynamic and structured categories. Tools must be developed for the integration of XQL functionality on top of the SQL in order to achieve complete integration of document and relational functionalities, based on a knowledge base containing different kinds of rules: -Visual rules, that characterize the different kinds of visual symbols; -Data rules, that specify the characteristics of the data model, the database schema, and the database instances; -Mapping rules, that specify the link between data and visual elements; Methodologies and tools for the support of cross-language retrieval from the nodes of the federated architecture must be developed.

Innovation principles in I.DB.I. Services

I. Device independent access to a world-wide digital repository of the Icon Arts in any form of religious bound art including interconnected publications and reports.

II. Service's availability 24h/day, 7d/week.

III. Reduced cost and difficulty of the icon (plastic art) content distribution thanks to the Internet.

IV. Possibility to merge knowledge thematically in one personalized book by making a selection among the large offer of digital content (articles, chapters, extracts, etc.)

V. Optimized management of perishable contents thanks to the maintenance service.

VI. Reusability, archiving and availability of the digital content in the "memory of knowledge" because the authoring studio of I.DB.I. offers users all the required tools.

VII. Possibility to create and expand links between small and specialized communities of East Christian icon arts researchers.

VIII. Technological efficiency for querying, reaching and online properly displaying retrieved multimedia content.

IX. Network performance monitoring of channel efficiency, channel propagation capabilities, time delay etc.

Socio-economic Innovation principles

It will deal with issues of co-operation within the context of an information society comprised of independent organisations with different rules, traditions, organisation structures, motivations for profit, nationality, laws, culture and languages. Potentially, it will have to integrate with people and organisation from all the continents of the world. Best practice methodologies and tools will be developed to meet the demands of this world-scale society.

INNOVATIVE SCHEME OF THE I.DB.I. DIGITYAL LIBRARY MANAGEMENT SYSTEM

We see the I.DB.I. DL MS architecture as an implementation of a next-generation Digital Library management system [1]. It should combine text and audio-visual searching, with new information visualization and relevantly personalized browsing tools, innovative interfaces, allowing retrieved information to be integrated, annotated and processed [5]. Processes are DL applications that derive from the specialized services available in the integrated system[2].

A case study of a next-generation Digital Library management systems found a prototype [6] matching very close the described innovative requirement to be developed under the DELOS[1] IST-FP5 EU project combining a rich set of features not available in this combination in any existing DL MS. Observing the integration activities of the second phase of the project.

The proposed for the middleware environment for P2P process execution technology is based on the Open Service Infrastructure for Reliable and Integrated Process Support (OSIRIS - developed at the ETH Zurich) as its base layer. Process design is largely invincible of details regarding services, including their physical location and availability which makes OSIRIS extremely reliable and scalable[2], compared to other service composition and process execution environments, incorporating sophisticated load balancing mechanisms without a centralized execution engine/scheduler.

On top of OSIRIS we propose to operate the Interactive SImilarity Search service layer (ISIS [2]) developed also at the ETH Zurich with the capability of mixing textual (metadata-based) and content-based queries with high performance indexing and retrieval in multimedia collections.

The Human-Computer Interaction Group at the Department of Computer and Information Science of the University of Konstanz has developed a visual information seeking system integrating novel interaction and visualization techniques in a consistent user interface design called Medio Vis [7]. It supports analytical queries as well as browsing-

oriented and opportunistic retrieval strategies. The integration of Medio-Vis in the DL MS offers the user new feature-based similarity search options supporting natural information seeking strategies.
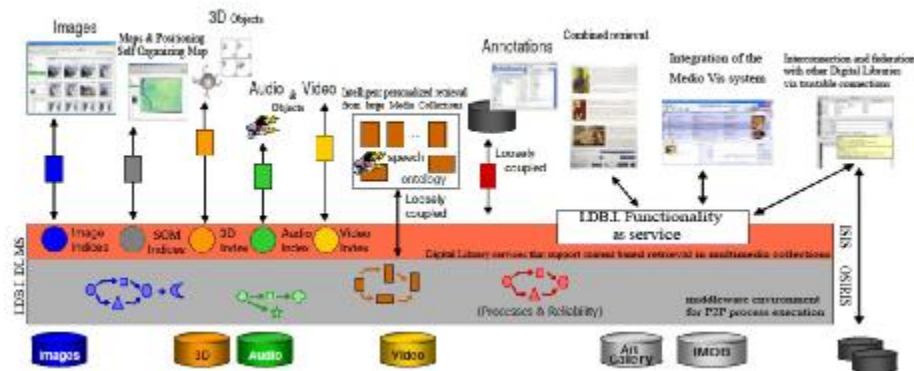
Fig. 2. I.DB.I. DL MS Overview

An important part of this architecture is flexibility and interoperability of the design making possible the on demand step by step modular integration of enhanced subsystems (layers over Osiris & ISIS) delivering services required by the users of this environment.

Network monitoring innovation principles

In a widely spread and distributed environment as such network provides, performance monitoring is a key issue. Having disparate network domains with different monitoring policy and support is putting difficulties in the process of problem resolution with the absence of common tools and independent evaluation. Such problems have been addressed in a special international group of network professionals for the EU in the frame of GEANT project, who have developed a universal tool for multiple network metrics called PerfSONAR.

Most appropriate for the current environment would be the overall view of important network interfaces using the Round Robin Database Measurement Archive (RRD-MA) service [10] functionality, which is operated through the unified interface of the PerfSONAR tool. RRD-MA service is providing and aggregating the statistics information on all utilization counters of all network interfaces registered in the system. The main network parameters being monitored are interface utilization, unicast and non-unicast packets and input/output errors. The most signifi-

cant benefit of RRD-MA deployment is the universal and centrally managed access to the instrument. Thus any network operator is able to view the statistics of any registered interface and what is more she would be able to quickly follow and focus on a special subset, if the traceroute field is used for a limited, filtered view. RRD-MA service is based on the SNMP collection of statistics, and is able to cooperate with multiple RRD based instruments as Cacti, Cricket and others.
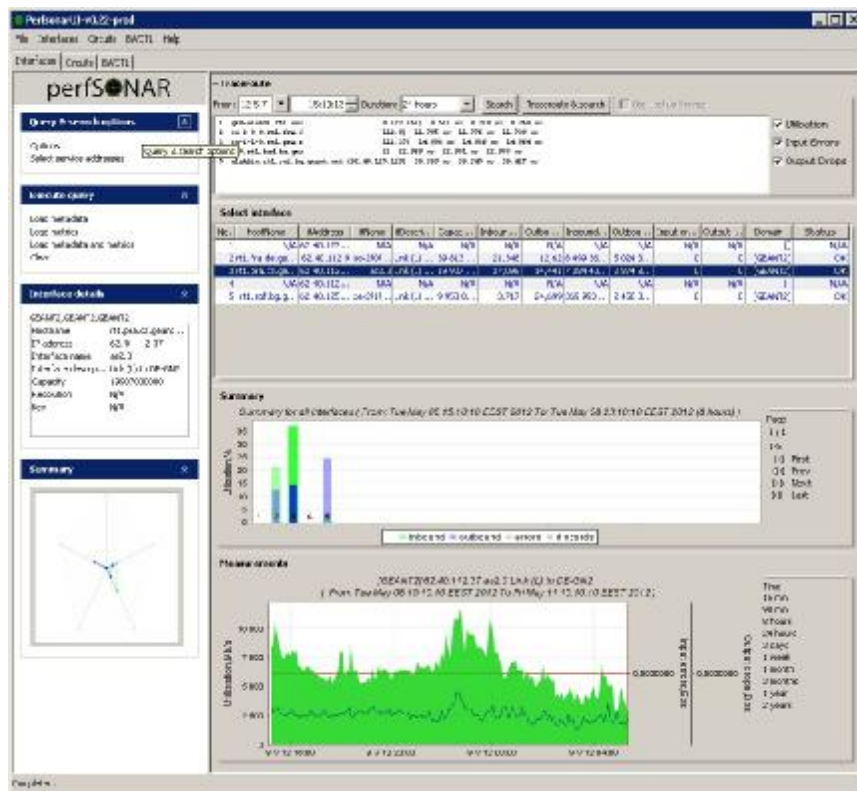
Fig. 3. PerfSONAR tool at the RRD-MA view with a focus on selected interfaces

DELOS DLMS with integrated subsystems – model for future extension of I.DB.I. DLMS

Having in mind the research done in the frame of DELOS project and the realized pilot DL MS with integrated agent based subsystems for search and retrieval e.g. DARE[6], Daffodil[6], Minerva[3], CoCoMa[7] and OntoNL[7] we may consider the next steps of evolution of the I.DB.I.

DLMS toward friendly user intuitive interface, flexibility, speed, federation, publication (iPaper [6]), charging, processing and analyzing the information and ensuring device independent access, taking into account the Intellectual Property Rights. Last research efforts led to the extension of the OSIRIS middleware which will allow making use of DelosDLMS and its services on mobile devices.
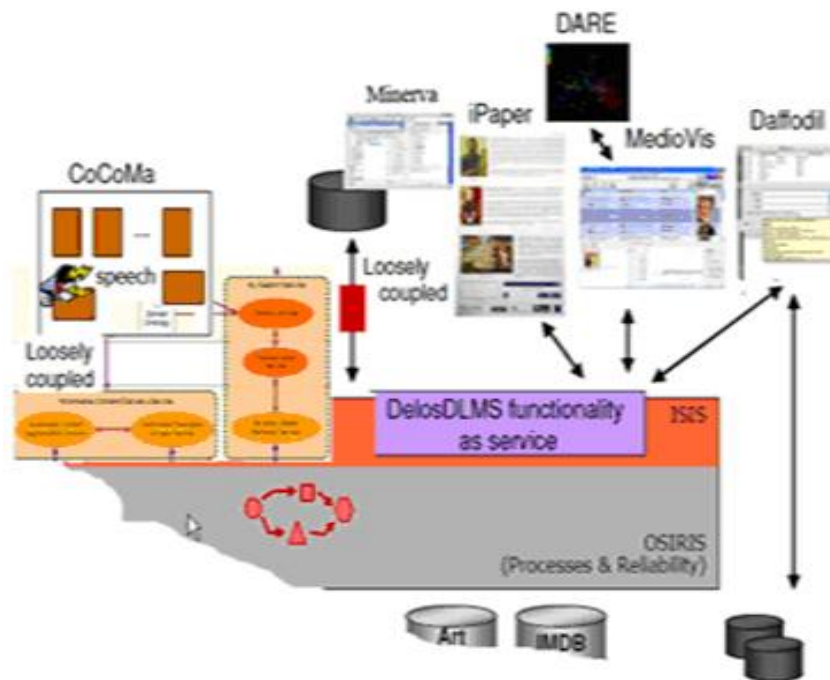
Fig. 4. Part of DELOS DLMS with integrated Subsystems

DARE – Drawing Adequate Representations, a system implementing a general theory for establishing the adequacy of a visual representation, once specified the database characteristics.

CoCoMa – subsystem, supporting complete multimedia retrieval scenario (search, selection, real-time content segmentation and delivery)

OntoNL – a natural language interface generator framework enabling natural language interfaces on top of knowledge repositories

MINERVA - is a fully operational P2P search engine which maintains a metadata directory that holds very compact, aggregated summaries of the peers' local indexes. The integration of the MINERVA P2P

search engine allows the execution of textual and content based queries over collections distributed in large scale systems.

Daffodil - a system of distributed agents for user-friendly access to Digital Libraries. Daffodil integrates several DLs freely accessible through the Internet to present to the user a single virtual DL.

CONCLUSION

So the way of bringing up to date the I.DB.I. Artery is moving it to I.DB.I. Digital Library in means of technology, services, Digital Library Management System and integrating all in one. The final goal is to develop a prototype for a next generation Digital Library Management System that combines a rich set of specific features matching the requirements set by the East Christian icon and plastic arts repositories and user requirements.

References
1. P A. Rowstron and P. Druschel 2001. Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems. In Middleware.
2. Schek H.-J. and Schuldt H 2006. DelosDLMS - Infrastructure for the Next Generation of Digital Library Management Systems. In ERCIM News, Special Issue on European Digital Library 66: 22-24.
3. M. Bender, S. Michel, P. Triantafillou, and G.Weikum 2005. Minerva: Collaborative p2p search. In VLDB
4. C. Grün, J. Gerken, H. Jetter, W. König and H. Reiterer 2005. MedioVis - A User-Centred Library Metadata Browser. In Proceedings of the 9th European Conference, ECDL, Research and Advanced Technology for Digital Libraries, Wien, Springer, 174-185.
5. B. Bustos, D. Keim, D. Saupe, and T. Schreck 2007. Content-based 3d object retrieval. IEEE ComputerGraphics & Applications, special issue on 3D documents, Jul/Aug 2007
6. M. Agosti, S. Berretti, G. Brettlecker, A. del Bimbo, N. Ferro, N. Fuhr, D. Keim, C.-P. Klas, T. Lidy, D. Milano, M. Norrie, P. Ranaldi, A. Rauber, H-J. Schek, T. Schreck, H. Schuldt, B. Signer, M. Springmann 2007. DelosDLMS - the Integrated DELOS Digital Library Management System. In: Post-Proceedings of the 1st DELOS Conference on Digital Libraries, Tirrenia, Italy. Springer LNCS.
7. Ceri Binding, Gert Brettlecker, Tiziana Catarci, Stavros Christodoulakis, Tom Crecelius, Nektarios Gioldasis, Hans-Christian Jetter, Mouna Kacimi, Diego Milano, Paola Ranaldi, Harald Reiterer, Giuseppe Santucci, Hans-Jörg Schek, Heiko Schuldt, Douglas Tudhope, Gerhard Weikum. DelosDLMS: Infrastructure and Services for Future Digital Library Systems
8. Metadata Encoding and Transmission Standard (METS) Official Website, http://www.loc.gov/standards/mets/
9. Virtual encyclopaedia of the Bulgarian iconography http://bidl.cc.bas.bg/

10. Ulisses Alonso (RedIris), Guilherme Fernandes (RNP), Maciej Glowiak (PSNC), Andreas Hanemann (DFN), Vedrin Jeliazkov (ISTF) , Uros Juvan (Arnes), S. Kraft (DFN), Loukik Kudarimoti (Dante), Roman Lapacz (PSNC), Danijel Matek (CARNet/SRCE), Michalis Michael (Cynet), Valentin Rosic (CARNet), N. Simar (DANTE), Hans Trompert (SURFnet), Sven Ubik (Cesnet), Verena Venus (DFN), Stijn Verstichel (Belnet), "Deliverable DJ1.3.3: JRA1 Phase III Implementation Report", July 2007
11. Pavlova-Draganova, L., Paneva-Marinova, D., Pavlov, R. (2011), Ontological Presentation of East-Christian Iconographical Art Domain, International Journal "Serdica Journal of Computing"

Author: Laboratory of Telematics to the Bulgarian Academy of Sciences –
1113 Sofia, G. Bonchev 8, Bulgaria, yoshinov@cc.bas.bg, i.garnizov@cc.bas.bg

INFORMATICS IN THE PUBLIC ADMINISTRATION, LAW AND CORPORATIVE SECURITY

ELECTRONIC COMMUNICATION WITH THE PUBLIC ADMINISTRATION

Eddie Chakarov, Ivo Panov

Abstract: The report presents a model of electronic communication between one application running in the offices of the judicial officers and private bailiffs and other WEB application of the National Revenue Agency. The communication is realized by extracting the data from the client-server application and automatically filling them in the WEB forms of the NRA. The most important benefits of the application are: Effective electronic communication between the national and the local registers of various state, judicial and public institutions; Structuring and automating the data transfer and in result reducing the number of the unintentional technical mistakes of the users; Acceleration of the information exchange between institutions; Increase of the efficiency of the public and private enforcements.

Key words: electronic communication, web application, judicial officers, JES

1. Necessity

The Electronic Government is based on efficient, accurate and reliable communication between the state, the judicial and the public institutions and the citizens.

The simplest type of such communication is the exchange of secure and reliable electronic messages signed with some type of electronic signature. Examples of such communication are the VAT declarations submitted each month to the NRA by all the entities registered for VAT in the country. Despite the obvious advantages, one disadvantage of this type of communication is the duplication of inputting the same data by various components in the process.

Another way of electronic communication is the data exchange between end applications running in different structures. A relevant example is connecting the information systems of the Customs and the Revenue Agency. As we know, building such "connectivity" is an extremely expensive and time consuming process in which some

technical but predominantly organizational problems should be resolved and dealt with.

A third way is by WEB forms which are being used more and more by governmental and judicial institutions for communication with citizens and businesses. The aim and advantage of these applications is to transfer to the primary source the input of the initial data of the public administration information systems. The disadvantage, as with the first type of communication, is the duplication of the completed information in the WEB form and in the forms of the local administration, which may exist in the office of a businessman.

2. Implementation

The paper presents a model of communication in which information is retrieved from a database maintained by a client-server application and is automatically filled into the WEB form issued by the NRA. The client-server application is the software system "JES" which automates record-keeping activities in the offices of state and private enforcement.

Several programming problems relating to the completion of the NRA WEB application forms have been resolved successfully.

Causing various events and recording data in different destination types in the web form is automated by the software. The first destination type is text fields, which completion is most obvious. Other field types include date, (which is selected from a calendar in the original application), radio buttons, dropdown menus and more. A further difficulty is that the forms submitted to the NRA must be signed with a universal electronic signature.

All these problems are successfully dealt with by creating the application – a specific type of browser that functions simultaneously as a standard browser and besides that has access to the data of the client-server application and utilizes the knowledge of the data structure of the software system " JES".

3. Electronic Enquiries in NRA

Following is the description of how the specialized software functions in order to enable the communication between the public and private enforcement and the NRA.

3.1. Making enquiries in "JES"
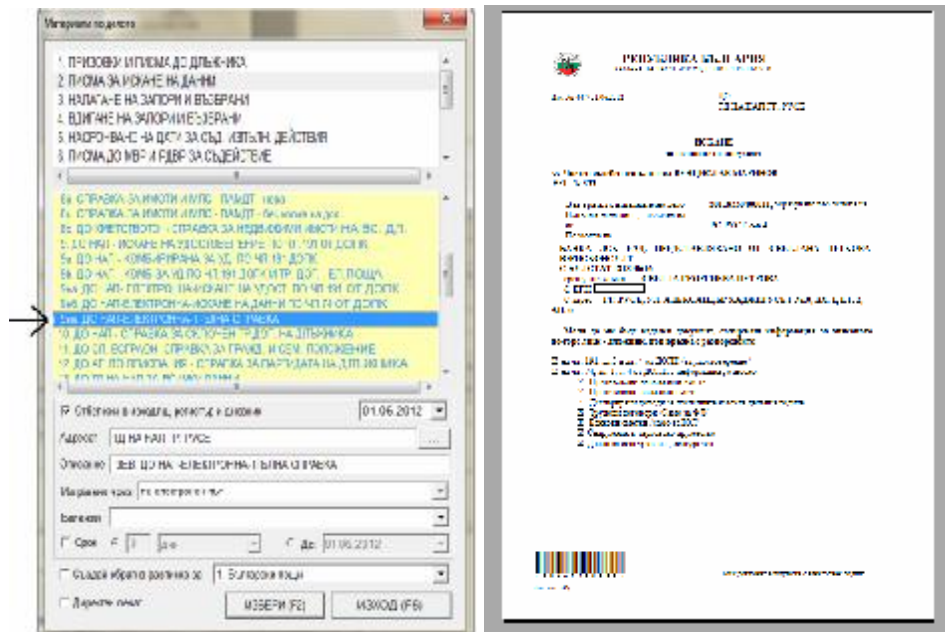
3.2. Transfer of Management to the Browser
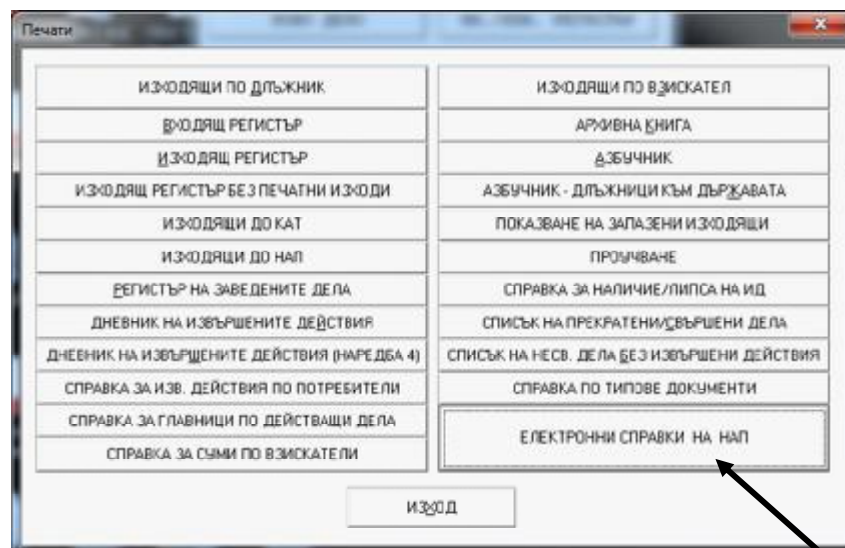
3.3. Navigating through the menus of the WEB application, authorization by electronic signature and opening empty WEB form
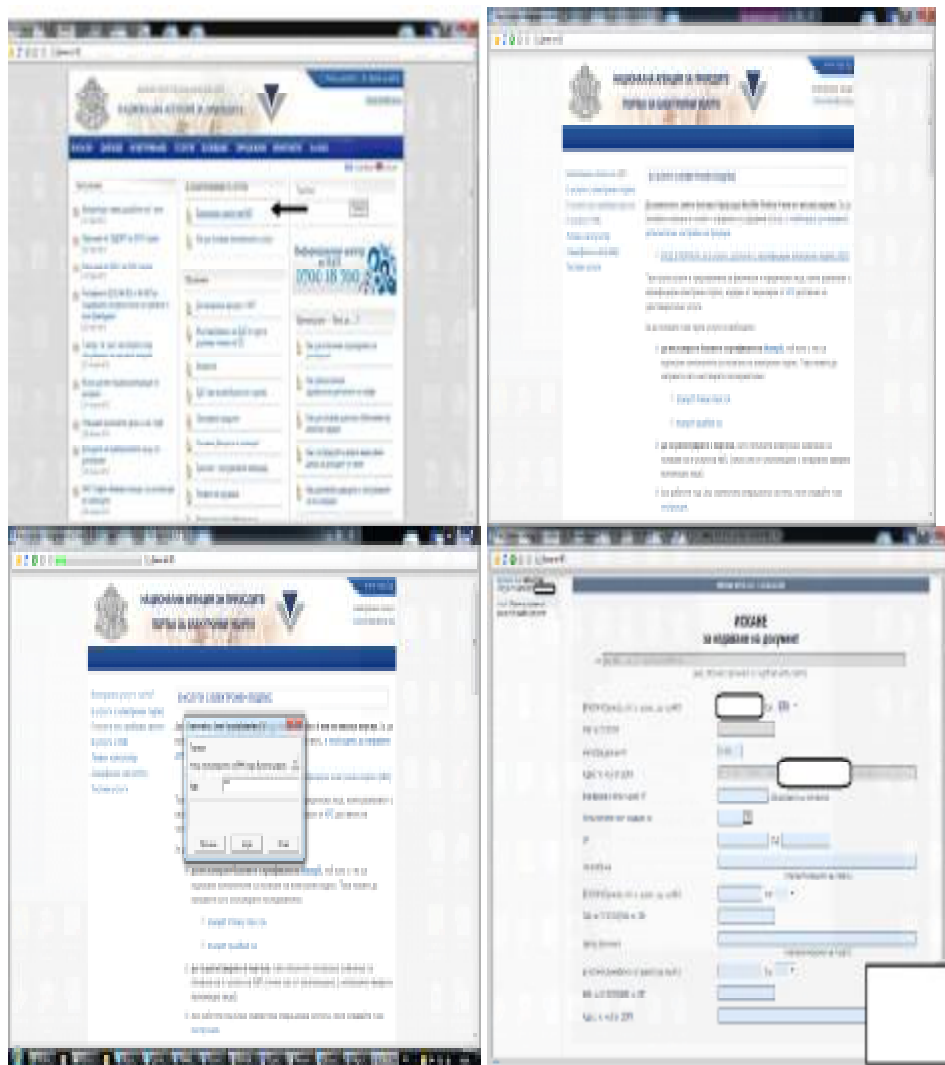
3.4. Retrieval of data from the client-server application and completing the WEB form

3.5 Signing the electronic form, submitting the form to the NRA and getting the receipt

3.6. Registering the issued document in the client-server application

In conclusion we would like to emphasize once again that the software discussed in the paper is already being used and is functioning well. It is applied in the offices and bureaus of over 150 public and private law enforcement executive magistrates and the daily exchange of information involving thousands of documents is carried out.

Authors:
Ass.prof. PhD. **Edi Yosif Chakarov** – Varna Free University "Chernorizets Hrabar", +359888973769, eddy@abv.bg
Ivo Evgeniev Panov, Master of Mathematics, GERION DZZD

ACADEMIC FREEDOM AND AUTONOMY OF INSTITUTIONS OF HIGHER EDUCATION IN REPUBLIC OF KAZAKHSTAN

Shynybekov D. A., Muhamedyev R. I., Abzhalyeva A. T., Hajiyev F. A.

Abstract: Questions of academic freedom and academic autonomy in the different countries - the USA, Russia, Kazakhstan, Germany are affected. The results of students and teachers polls are analyzed. The issues of transition of higher educational institutions in Kazakhstan to a greater degree of academic freedom and autonomy are analyzed.

Key words: Academic freedom, academic autonomy.

Academic freedom and academic autonomy

The basic principles of academic freedom and academic autonomy adopted in 1940 by the American Association of University Professors, although the principles enunciated at the conference in 1925 [1]. According to these principles, "The teaching staff in the course of their duties shall have freedom to express their views on the subject".

The Administration may impose "restrictions on academic freedom for religious and other issues" that must be clearly specified in writing at the time of employment. Academic freedom in almost all universities in the U.S. is perceived as an asset of scientific world. In private universities item about academic freedom protection of professors is usually included in his contract or statute of the University [6]. At present these principles in general terms include [2]:

• Both inside and outside the educational institution or research organization is allowed full freedom to put any questions and look to truth, including on the controversial and unpopular opinions, whether or not to hurt anyone or that point of view.

• Educational institutions and research organizations have no right to restrict academic freedom for its staff, and use their public statements as a reason for disciplinary action or dismissal.

The concept of academic freedom are closely linked with the concept of tenure (long-term or permanent hiring of lead professors). As noted in [4] "Academic culture is based on the principle of academic merit (meritocracy), the freedom of research initiatives and on the principle of competition, complemented by some elements of cooperation and mobility". This is the basis of world-class universities. Many universities in the United States and Canada provide the elected representatives of the senior faculty members holding positions Professor and Associate

Professor, lifetime contract (tenure), without the right of dismissal by the administration. The overall objective of the contract - to protect the independence of scientists (academic freedom). Lifetime contract with the university provides an incentive to work on the reputation of this institution without fear of competition from talented colleagues and engage in research in the "unfashionable" areas of science. In addition, this system protects the specialist in a tough market and authoritarian state system. A scientist can work without paying attention to the challenges of changing market demands and conjuncture government agencies or individual officials.

The European experience of academic freedom and autonomy

In recommendations of UNESKO «Academic freedom requires the autonomy of institutions of higher education. The autonomy of management is needed for effective decision making, and in area of academic life, standards of management. Academic freedom should be welcomed. Autonomy is form of academic freedom of staff to have a higher education institution empowered to take part in the government, but also have the right to choose the appropriate representatives".

Thus, universities should have the right to:
• Allocate financial resources
• Create their own development strategies;
• Select priorities for educational and scientific work;
• Create your own training programs;
• Establish criteria for the selection of teachers and admission of students;

Despite the general commitment to established principles, their implementation in different countries is very different [2]. Right to life-long occupation of positions in the 1980s was limited in the UK, it is not in Australia and many European countries. In addition, most European universities for graduate students and young researchers, regardless of the presence or absence of the teaching degree is prohibited. In France, the lifetime employment - common, even for young scientists. In New Zealand and certain other countries practice similar in the sense of long-term rentals. In Russia, a system of contractual employment, the revision of the employment contract every 3-5 years, which in many circumstances, makes it possible to lifetime employment. Nevertheless, conditions of employment of faculty and academic staff members are not guaranteed, for example, dismissal on grounds of redundancy. German universities have also adopted the general principles of academic freedom is defined, including the Bologna Convention. As in the U.S., a

120

candidate for a professor to be recognized outside of the University and, for example, may be invited to another University before receiving his professorial positions at the University. Allowed in exceptional cases, employment professionals professorial positions without any academic degree.

The level of guarantees of word freedom and academic freedom in Europe at the legislative level, illustrated in Table 1 [4]

Table 1. Comparison of aspects of academic freedom and autonomy in Europe

Nation	Constitutional protection	Specific legislative protection	Self-governance	Appointment of rector	Academic tenure
Finland	High	High	High	High	High
Slovenia	High	High	High	High	High
Czech Republic	High	High	High	High	High
Hungary	High	High	High	High	High
Spain	High	High	High	High	High
Latvia	High	High	High	High	Medium
Lithuania	High	High	High	High	Medium
Slovakia	High	High	High	High	Medium
Poland	High	Medium	High	High	High
Austria	High	Medium	High	High	Medium
France	Medium	High	Medium	High	High
Portugal	High	Medium	Medium	High	High
Italy	High	Low	High	High	Medium
Estonia	High	Medium	High	Medium	Medium
Greece	Medium	N/A	Medium	High	High
Germany	Medium	High	High	Medium	Medium
Ireland	Medium	High	Low	Medium	High
Luxembourg	Medium	Medium	Medium	High	Medium
Sweden	Medium	Low	Low	Medium	High
Malta	Medium	N/A	Medium	High	Low
Denmark	Medium	Medium	Low	Low	Low
Netherlands	Low	Low	Low	Medium	Medium
UK	N/A	Low	Low	Low	Low

Asian experience in higher education

In [5] noted that Asia has a strong high-quality higher education sector, for example,

• Singapore and Hong Kong have achieved considerable success by simply copying the Western model.

• South Korea financed holding of several national companies to develop the academic system

• Taiwan has emphasized the conviction of his former compatriots who received Western education, to return to his homeland for the development of national universities.

Kazakhstan also relies on the involvement of foreign experts on the one hand, on the other side of their own finances training of personnel in leading foreign universities.

However, a number of structural, academic, and cultural factors could keep even the best Asian universities achieve peaks of academic quality, which include [5]:

• The strong influence of informal relations ("Guangxi" chinese), as a consequence, an autonomous "self growing" faculty staff, which makes innovative development.

• Linking career development with family ties, which is contrary to the principles of academic performance and academic freedom.

• Adherence to traditional methods of teaching and research that is not conducive to the development of independent thinking.

• The significant impact of the hierarchy, based, for example, on positions and respect for age. It is hard to run informal communication with teachers and leaders.

• The academic corruption and plagiarism.

• The complexities of the internationalization of universities.

• Low salary and large amount of load of teachers. The volume of workload is often so high that the research work becomes inaccessible for teachers.

• The lack of guarantees of permanent academic positions, ensuring academic freedom. Vulnerability of trainers workplace.

Thus, teachers feel need for greater protection of workplace and need for more funding, while at the same time, combined with sufficient competitive environment to encourage high performance.

Kazakhstan movement in the direction of academic freedom and autonomy of universities

Kazakhstan as well as other countries move in the direction of providing more freedom and autonomy of higher education institutions. The general vector of development has set by state program. It may be noted that

• Has adopted State program, which provides that guidelines of institutional autonomy will be developed – autonomy in the imple-mentation of educational, scientific, financial, international and other activities based on Nazarbayev University model.

• Since 2015 will be presented autonomy for national research universities since 2016 – for national universities, since 2018 - the rest of universities.

- The Law of Republic of Kazakhstan "On the status of" Nazarbayev University" has adopted. It provides that
  o The University is an experimental platform, which perform developing, monitoring, research, analysis, testing, deployment and implementation of innovative programs in education and science.
  o The right of University to carry out educational activities without a license and state certification.
  o Independence of admission of foreign workers.
  o Independence in the development of educational programs and organizing of educational activities

Teachers attitude to principles of academic freedom in Kazakhstan

To investigate the views of faculty members from various universities were interviewed. In the survey attended by 65 teachers IITU, UIB, KazNU, MAB, KazNTU Satpaeva n., AUE&S, CAU. Surveys showed that
- Important in the definition of "academic freedom" teachers of all schools, regardless of age, ownership of the university, say - to define their own strategies, choice of priorities in teaching and research and development of educational programs.
- About necessary of State program, where will develop the principles of autonomy of universities, the responses were divided:
  o Teachers of private and public high school institutions were undecided (23%) or think that principles of this autonomy must develop university himself. In particular, they suggest that it is part of academic freedom.
  o Teachers of public high schools, particularly universities, which have status of "national", expressed the view that such a program is needed (86%).
- Teachers in the vast majority (91%) believe that it is necessary to depart from the concept of "national standard of education" and "typical program."
- Many (83%) think that with the above measure each university should "independently to determine the order of educational activities, admissions for training, monitoring of academic progress, interim and final assessment", "to carry out educational activities based on self-designed training programs" and "provide opportunities to develop author courses".
- 99,9% of respondents believe that Kazakhstan education system should not copy the Western model.

- Despite the fact that in many universities around the world believed that lifetime contract with the university provides motivation to work for reputation of the school, the respondents think that it must not enable a right to lifetime employment positions in the universities of Kazakhstan. The reasons for this, by most accounts, the following: "People will calm down and stop growing", "should be given an opportunity for young teachers to express themselves, and for this they should at least get to college", "change management", etc.

- 97% of teachers consider that constraining factors in the development of Kazakhstan's education system is "a lot of the load of teachers, in which condition the research is inaccessible to teachers." 82% of workers in higher education believe that "teaching and research carried out by conventional methods," which does not contribute to the development of independent thinking, and 80% - "low salary". Also was mentioned such answers as "academic corruption and plagiarism" - 17%, "the relationship career with family ties, which is contrary to the principles of academic performance and academic freedom" - 15%, "strongly influenced by informal relations" - 12%, "the hierarchy, based on the positions and respect for age "- 8%. In this issue we could find some answers.

- 94% of teachers consider that only the university must authorize its own training program and, in any case the Ministry of Education and Science or any other specially authorized must not do it.

The teachers are consentaneous in another question regarding the maintenance of the course. The results showed that prior to the course instructor is obliged to prepare a syllabus only. Many respondents argued that one course does not require a large number of documents - UMCD, work programs and so on.

In the matter of courses and professors choice opinions were divided. All consider that choice is necessary. But, more than half (56%) think that the choice should be started only from the 3rd year because on the 1st course the student is not ready to choose. The second half of respondents (44%) think that students can choose teachers and discipline from the start of study because it is not just the right but proper and necessary element of education system. In addition all teachers think it is absolutely not necessary to have a special schedule of presence in the workplace at the Department. And also should specify the time of presence, in addition to the training sessions and consultations.

All teachers without exception believe that academic freedom should have limits, though not specified areas.

Conclusion

Academic freedom as an opportunity to express their views and to form its own rules of educational institution is an important feature world-class universities.

Kazakhstan higher education inherent characteristic of many Asian universities of problems that can significantly slow further development. At present, Kazakhstan is only one university is fully consistent with the principles of academic freedom and autonomy. At the same time, teachers of universities in Kazakhstan as a whole are ready to adopt the principles of autonomy and freedom as defined in the recommendations UNESKO. A number of legislative measures contribute to the movement of Kazakhstan in this area. In this process, obviously, will be worked out in some way different from the conventional way of implementing the principles of academic freedom and autonomy.

References
1. 1940 Statement of Principles on Academic Freedom and Tenure.American Association of University Professors (AAUP) http://www.aaup.org/AAUP/ pubsres/policydocs/contents/1940statement.htm
2. Wikipedia Академические свободы http://ru.wikipedia.org/wiki/Академиче-ские_свободы
3. Terence Karran. «Academic Freedom in Europe: A Preliminary Comparative Analysis». Centre for Educational Research and Development, University of Lincoln, Lincoln LN6 7TS UK.
4. Сколково: культура и этика. http://www.ru.biocenter.kz/scienceandeducation/ 901-skolkovo-kultura-i-etika-.html
5. Филип Г. Альтбах. Век азиатского высшего образования? http://ihe. nkaoko.kz/archive/ihe59/2
6. Льюис М. Бенедикт Академическая свобода в США на распутье. http://ihe.nkaoko.kz/archive/ihe57/3
7. Серж Хейвард: Академические свободы и нравственные нормы станут ядром проектов «Сколково». http://skolkovo-ru.livejournal.com/149360. html?thread=410480

Authors:
Shynybekov D.A., professor, rector IITU, Almaty
Muhamedyev R.I., professor, head of chair, IT department, IITU, Almaty
Abzhalieva A.T., docent, IITU, Almaty
Hajiyev F.A., doctor, Nazarbayev University, Astana

Contacts:
R.I. Muhamedyev
International IT University, IT department,
Manas/Jandosov Str., 34A/8A Almaty, Kazakhstan,
e-mail: ravil.muhamedyev@gmail.com
phones: +7 701 3393720,+7 777 2399788

DESCRIPTORS OF UNCERTAINTY AND UNCLARITY OF RISK OF ENVIRONMENTAL DANGER ACTIVITIES IN BORDER'S ZONES

**Lyubomir Vladimirov, Nikolai Kovachev,
Plamen Manev, Ventsislav Dobrinov**

***Abstract:** The insecurity in the measuring is a problem which the researchers of the risk avoid. A taxonomy which includes categories, classes, groups and sub-groups is proposed in order to be revealed the reasons for the occurrence of the insecurity in the present work. They are defined and interpreted. Expert and mathematical methods for assessment are available. The possibilities for establishment, limitation and reduction of the insecurity are brought out.*
Key words: *information, insecurity, assessment, risk.*

The main disadvantage of the risk's researches from trans-border ecologically dangerous activities is included in the ignorance of the information uncertainty in risk design, analysis and assessment. This influences the objectivity and causes undertaking non – argument defense solutions. That's why it is necessary to develop the problem with the uncertainty methodologically, which will the base for rules to reduce the risks and requirements system creation. They could be part of the presented in [1] idea for trans-border ecological gravities.

The object of the present paper is defining the descriptors of the uncertainty and unclarity of risk of environmental danger activities in border's zones.

The main purposes are:

1) Structure creation and descriptor's definition of the uncertainty and unclarity;

2) Method for uncertainty qualitative analysis and assessment proposing;

For the first problem solution, as a methodological base for the analysis, the definition of S. Dentchev and D. Hristozov [3] for information medium uncertainty is acquired. In the present paper this definition is developed with a new structure. The definitions and interpretations of the main uncertainty's descriptors in risk-classes assessment, categories, groups and elements are included. The two types of uncertainties, defined by S. Dentchev and D. Hristozov [3] are used – Uncertainty and Unclarity.

Our experience [1,2,6,7] presents that the uncertainty can be described qualitative using two categories of descriptors:

I category. Information's diversity descriptors - Narrow restricted, reflecting the spectrum of the risk information values. It is defined by:

1) Fortuity/ Constantly of the values and the indications for risk formalization,

2) Mutability of the values,

3) Dispersing of the indications values, defining the risk levels,
4) Permissible in dangers structuring,
5) Approximation in risk information processing,
6) Methods for information processing,
7) Capacity of the data excerpt;
8) Software;
9) Select Distribution of the statistical low's distribution,
10) Error Measuring.
II category. Formalization risk descriptors - Risk Punctilious:
1) Danger Structure;
2) Space Situation;
3) Evens Chronology ;
4) Phenomennons/Impacts orderly;
5) Situations Metric;
6) Situations Fixing;
7) Cause/Effect Relation;
8) Vulnerability;
9) Context;
10) Injury;
11) Scenarios Logic;
12) Scenarios Structure;
13) Scenarios Function;
14) Criticality;
15) Environmental Insecurity;
16) Environmental Security;
17) Scenarios Selection;
18) Cascade Effect.
Unclarity is defined by two descriptors categories:
I category. Subjective unclarity descriptor - Subject Unclarity:
1) Risk sensation;
2) Risk establishment;
3) Risk Differing;
4) Risk Identification;
5) Risk Assessment;
6) Risk Estimation;
7) Risk Conclusion;
II category. Descriptors of Risk Linguistics:
1) Linguistics Unclarity, expressed in vocabulary using with impropriate terminology;
2) Finical Unclarity or ignorance of the rules for phraseology risk's description of risk situations and scenarios,
3) Semiotic Unclarity representing in the acquired, verified risk's and danger's structure plurals, syntactic and semantic rules and axiomatic, ignorance.
Risk Linguistics descriptors are base for language-logic memory used for definition's minds, reasoning, conclusions and assessment of

the risk, other persons communications and solution formulation. But for each element the language richness is used for linguistic reproduction. In other words, the linguistic risk model of the situations and phenomennons is formed.

Risk Linguistics in not caused from the subjective uncertainty, but is accompany and a necessary sphere. It is a result from the processes of risk's sensations, perceptions and an idea, a form for risk presentation and realization. On this type of problem, the researches of the risk from ecologically danger objects don't give an attention. They miss the fact that everything to discuss has its linguistic expression. Including the risk scenarios upon which are based lots of investigations [6]. Whey they are not defined in the term of language, they are not able to use, in spite of their structural objectiveness. Three types of reasons cause the uncertainty based on the linguistics: 1) ignorance of the specific lexical elements; 2) ignorance of the reasons and the resources for the cause and effects relationships; 3) ignorance of the risk semiotic.

The consequences from these reasons may be set when analyzing of the documents for the appeared critical phenomennons. Our experience in reasons for criticalities in some economics activities appearance investigations verifies that [4,5]. The documents from accidents, pollutions, fire, and other events are used such definitions and phrases, juxtapositions and descriptions, making the information unrecognizable. Consequently in order to limit the uncertainty of the linguistic risk models it is necessary to limit the reasons mentioned above and, if there it is possible, to remove them.

The shown descriptors allow systematizing the reasons and the character of the uncertainty and unclarity in risk definition. Besides the inserted classes, categories groups and elements, a systematization upon 3 additional indications – determinativeness, assessment and manageable is made.

The purposes, besides the uncertainty and unclarity's definition in their different variety, is to formulate the indentifitivness in order to fix the level and to assess the possibilities and the ways to remove, limit or reduce the uncertainty. This is the way to follow a logically verified consequence.

The ability to set represents the potential to reveal categories, groups and element of the two types uncertainty- indefiniteness and unclarity. The preliminarily hypothesis for it are checked logically ant their extent is defined. If an element of the uncertainty and unclarity could be defined it means the ability to set them.

The assessment is a qualitative representation of the type, essence and the extent of the unclarity and uncertainty. It gives possibilities for numerical definition of the importance of the different elements of the uncertainty from one side, and to analyze the uncertainty in different linguistic and juxtaposed to them poli- and monosituation risk models. Consequently the assessment is an indicator for model's authenticity.

The ability to manage shows the possibilities for influence upon the reasons and reduces the impact on the objectiveness of the risk assessment. It is a base for development of a system with methodological requirements for uncertainty limitation.

The importance of these three types of indications is shown axiomatically. They are the result from the accumulated experience, published in [1,2,3,4,5,6,7]. The analysis shows some tendencies:

I. All the descriptors of the uncertainty and uclarity are fixable. Some of them could be set using direct methods, and the other requires indirect methods, there is another part, uses combine method.

II. The descriptors could be assessed mainly by combined method – expert and numerical. The numerical methods are applicable for the elements, linked with the entropy of the phenomennons, impacts and effects, variations and fortuity.

III. The main part of the descriptors of the uncertainty and unclarity are controllable. The uncontrollable are linked with the natural occurrence and processes, with alternations and chances, including subjective character. Here is a moment that subjective uncontrollable descriptors cold be transformed in controllable by training, requirements for the behavior of the subjects, control of the knowledge about the risks, the specific person makings, relationships to the certainty and etc.

The uncertainty and unclarity's assessment could be made using two main methods – expert evaluation and mathematical assessment. Their division is under condition, because the result processing from the analyses using expert method is made numerically. The goal of the expert assessment is to fix the uncertainty and unclarity in linguistic and graphically described risk model.

There are two purposes for risk definition:

1) Definition of the importance of the uncertainty and unclarity elements and their graphic interpretation using poli- and mono situation graphs.

2) The extent of the uncertainty's elements for a given model and risk graph. The organization and the implementation of the expertise in methodological attitude are made according to the experience, requirements and indications in [6].

Taking into account the specification of the assessment, aim and purposes of the research, two criteria for assessment of the descriptors of the uncertainty are inserted - $Rank$ and $Rait$. Two levels of importance are introduces:

A) For the rank criteria $Rank$: importance (rank 1); clarity (rank 2); accessibility (rank 3); definiteness (rank 4);

B) For the $Rait$ criteria: very large (rait 1); large (rait 2); middle (rait 3); little (rait 4);

An excerpt from a form for assessment of the uncertainty is introduced in the table 1.

<div align="right">Table 1</div>

An excerpt from a form for uncertainty's assessment

Descriptor's designation	Descriptor's name	Descriptor's Rank				Descriptor's Rait				Rank	Rait
		importance (j_1=1)	clarity (j_2=2)	accessibility (j_3=3)	definiteness (j=4)	very large (j_1=4)	large (j_2=3)	middle (j_3=2)	little (j_4=1)	Rank	Rait
D1	Fortuity	c1.1	c1.2	c1.3	c1.4	d1.1	d1.2	d1.3	d1.4	Rank 1,	Rait 1,
D2	Mutability	c1.2	c2.2	c2.3	c2.4	d2.1	d2.2	d2.3	d2.4	Rank 2,	Rait 2,
D3	Dispersing	c1.3	c3.2	c3.3	c3.4	d3.1	d3.2	d3.3	d3.4	Rank 3,	Rait 3,
D4	Permissible	c1.4	c4.2	c4.3	c4.4	d4.1	d4.2	d4.3	d4.4	Rank 4,	Rait 4,
D5	Approximation	c1.5	c5.2	c5.3	c5.4	d5.1	d5.2	d5.3	d5.4	Rank 5,	Rait 5,
D6	Methods	c1.6	c6.2	c6.3	c6.4	d6.1	d6.2	d6.3	d6.4	Rank 6,	Rait 6,
D7	Capacity	c1.7	c7.2	c7.3	c7.4	d7.1	d7.2	d7.3	d7.4	Rank 7,	Rait 7,
D8	Software	c1.8	c8.2	c8.3	c8.4	d8.1	d8.2	d8.3	d8.4	Rank 8,	Rait 8,
D9	Select Distribution	c1.9	c9.2	c9.3	c9.4	d9.1	d9.2	d9.3	d9.4	Rank 9,	Rait 9,
D10	Error Measuring	c1.1	c10.	c10.	c10.	d10.	d10.	d10.	d10.	Rank 10	Rait 10
D11 ..	...	c1.1	c11.	c11.	c11.	d11.	d11.	d11.	d11.	Rank 1	Rait 11
D12 ..	...	c1.1	c12.	c12.	c12.	d12.	d12.	d12.	d12.	Rank 12	Rait 12
D13 ..	...	c1.1	c13.	c13.	c13.	d13.	d13.	d13.	d13.	Rank 13	Rait 13

The responses of the k expert according to rank criteria $Rank_k$ constitute the vector $c_{i,j}$ ($c_{1.1}$, $c_{1.2}$, ….$c_{4.20}$). To define the level of $Rank_k$ the ranging method is applied and the dependency is used:

$$Rank_{ik} = \frac{min_k \sum\limits_{k} c_{i,j_k}}{\sum\limits_{k} c_{i,j_k}}, \qquad (1)$$

where c_{ij_k} is the rank of i descriptor on the j value of the rank, evaluated from the k expert. The shown by the k-th expert values of the rait criteria $Rait_k$ form the vector $d_{i,j}$ ($d_{1.1}$, $d_{1.2}$, ...$d_{4.20}$). For their value definition the method for numerical assessment with dependency are applied:

$$Rait_{ik} = \frac{\sum_k P_{ij_k}}{\sum_i \sum_j P_{ij_k}} , \qquad (2)$$

where $P_{ij_k} = d_{ij_k} / \sum_j d_{ij_k}$; and d_{ij_k} is the assessment of the k–th expert upon i–th descriptor and j–th value. The impotencies of the criteria $Rank_{ik}$ are used for importance definition of the presented descriptors of the uncertainty, revealing their importance extent when fix the common uncertainty in the environmental danger objects. The criteria $Rait_{ik}$ is applied for the extent of insecurity's definition of the risk's models using descriptors. Both the criteria must be used for each variety or branch of the models definition. The assessment accuracy depends on the number and the values of both the criteria. The above shown divisions are adopted on the base of the requirements, given in [6]. The scales could be enlarged, inserting intermediate values, but they increase the responses vector, complicated their usage.

The numerical assessment aims to set the value of uncertainty for risk phenomennons, impacts and effects, described with quantitative formulated indications. It is linked mainly with the variety, representing the information changes, nature and individual subjective heterogeneity, time, space, processes and conditions.

The ignorance of the uncertainty and unclarity as main insecurity characteristics leads to mistakes, inaccuracy, unverified claims, irregular assessments and measurements. In order to solve the problem it is necessary to determine the following purposes:

1) To set the type of the variables, used for risk design.

2) To choose methods for insecurity's presentation.

3) To fix numerically the degree of uncertainty when measure a particular designed risk.

For the first problem, it is necessary to analyze the presence or experimentally settable information and to estimate if the variables are occasionally. They must adopt only one from all the possible values that could not be predicted until the experiment is conducted. In this aspect it is necessary to accurate set the character of the random quantities – discrete or indiscrete, because it influences on the choice of the presentation form and the manner to represent the insecurity.

On the second and the third problems solution the uncertainty could be analyzed by:

1) The order and the distribution laws at discrete or indiscrete random quantities [6,7]

2) Analysis of the boundaries for the appearance probability [3,6,7];

3) Shannon's Entropy [3,6,7];

4) Hartley's information [3,7];

5) Boltzmann's Entropy [3,7];

6) Analytical methods: Taylor's decomposition [6,7] and approximation [6,7];

7) Method "Monte Carlo" for risk design and Latin hyper squares [3,6,7]

8) Interval arithmetic [6,7];

9) Measures for unclearness [3,6,7].

The first three methods are used in our investigations [1,2,4,5,7] as long as they are based in the distribution laws of the random quantities. For estimation of the uncertainty we use the mean value $\overline{X}$, dispersion $s^2[X]$, standard deviation $s[X]$ or variance $V[X]$. Additionally these quantities are analysed as random and the uncertainty's distribution laws are settled. In the presented analysis of the uncertainty we propose using the discrete laws distributions and 18 laws for indiscrete quantities. The hypothesis verification is made with the c^2, criteria, Anderson-Darling test $A-D$ and Colmogorov-Smirnov test $K-S$. The Risk 4.5 software, a product of Palisade Corporation is used.

Based on the essence of the risk assessment method [1] if follows to mark the importance of the boundaries to determine the probability for random quantity and the linked uncertainty. The envelopment mast be defined accurately, because of it importance upon risk results.

The arguments depend on the topic, purposes, envelopment and limitations of the investigation and the specification of the assessed object.

Shannon's Entropy is applied successfully in the researches of S. Dentchev, D. Hristozov and V. Tomov [3,6]:

$$H(P(x)) = -\sum_{x \in X} p(x) log_2(x), \qquad (3)$$

where $p(x)$ is a probability distribution upon the multitude X, whereas $x \in X$.

For comparative analysis of two multitudes X and Y the simple, combined and conditional entropy and information transmission could be used [3].

On the base of the presented materials we make the following conclusions:

The structure of uncertainty's and unclarity's descriptors is composed. It represents the essence and definitions of the composed classes, categories, groups and elements. A descriptor's analysis is made.

Expert and numerical assessments are introduced. They allow determination of the value of the uncertainty and unclarity and the importance of the base components to be performed.

This complements the risk models, they become completed and full appearance and allow accurate risk assessment of environmental danger activities in border's zones.

References
1. Владимиров, Л. Рискметрия в екологичната сигурност. Монография. Варна, Варненски свободен университет "Черноризец Храбър", 2009, с. 279.
2. Владимиров, Л. Теория на риска. Част I. Опасности и заплахи, рискове и критичности. Русе, Русенски университет, 2011. 270 с.
3. Денчев, С., Д. Христозов. Несигурност, сложност и информация: Анализ и развитие на несигурна информационна среда. Книга 2. София, Издателство "Захари Стоянов", Унив. изд. "Св. Климент Охридски", 2004, с. 179.
4. Томов, В., Л. Владимиров. Риск таксономия. Габрово, Технически университет, Международна научна конференция "Унитех 07", част II, 2007, II302-II 307.
5. Томов В., Л. Владимиров. Изследване на критични ситуации от запалване на материали и замърсяване на въздуха. Част I. Методика на изследването.Русе, Известия на Съюза на учените, серия "Технически науки", 2007, № 6, 87–94.
6. Томов, В. Теория на риска. Монография. Русе, Русенски университет, 2003, с. 440 .
7. Томов, В. Л. Владимиров, М. Тодорова. Информационна несигурност в оценката на риска от екологично-опасни обекти. Шумен, Национален военен университет „Васил Левски", Факултет "Артилерия, ПВО и КИС", Катедра „Информационна сигурност", Научна конференция "Проблеми на информационната сигурност през XXI век". Сборник научни трудове, 2011, с. 221–228.

Authors:
Assoc. prof. PhD Lyubomir Vladimirov Vladimirov
Chief Ass. Nikolai Iordanov Kovachev
Chief Ass. Plamen Mihailov Manev
Chief Ass. PhD Ventsislav Dobrinov
Rousse 7017, Str. „Studetnska" № 8, University of Ruse „Angel Kanchev,
E-mail: lvvladimirov@uni-ruse.bg, nkovachev@uni-ruse.bg, pmanev@uni-ruse.bg, vdobrinov@uni-ruse.bg

DANGERS, RISKS AND CRITYCALLITIES DESIGN IN THE ENVIRONMENT OF THE BORDER'S ZONES

Lyubomir Vladimirov, Nikolai Kovachev, Plamen Manev, Ventsislav Dobrinov

Abstract: Dangers, risks and caused criticalities at the environment, models are presented in the paper. Some mathematical methods and manners for numerical prediction are argumented. The limits for alternation are defined. The risks and criticalities are graded.

Key words: hazard, risk, criticality, environment, border zone.

The paper aims to bring out an analytical method of the border's environmental dangers, risk and criticalities. For the solution two purposes are solved: 1) Risk from appearance of border's dangers modeling, 2) Formalization of the model of the risk and criticalities, 3) Choosing of numerical method for quantitative modeling, 4) Border's fixation for criticalities variation, 5) Criticalities gradation.
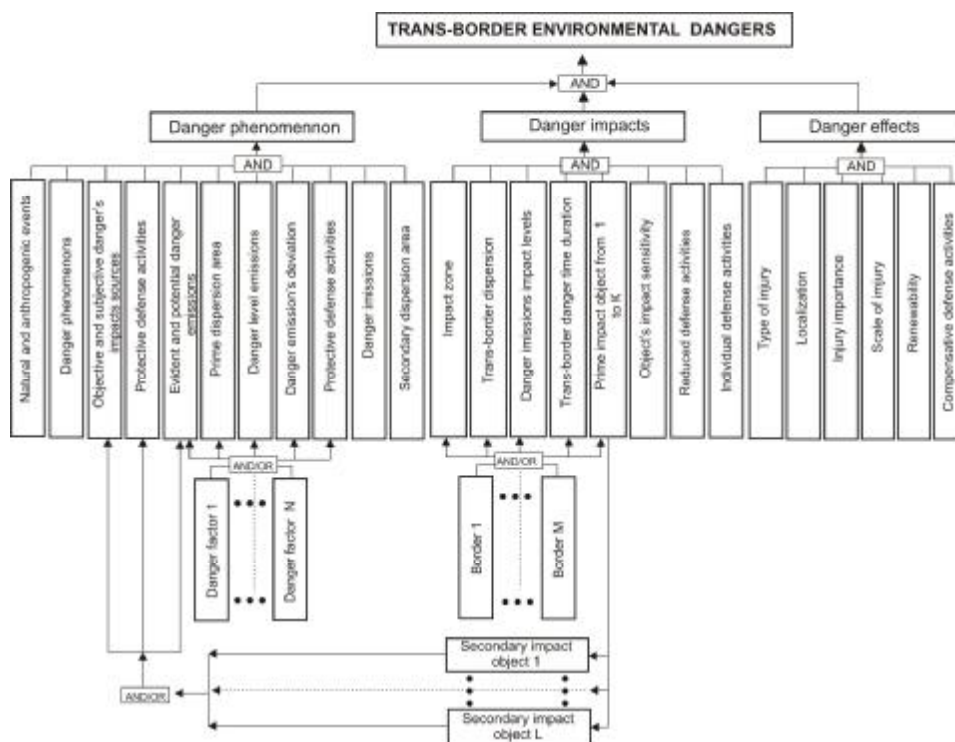
Fig. 1. Morphological border's dangers model

Risk and criticalities modeling is based on the model of the trans-border danger, shown on fig 1 [5].

Our experience shows to us [1,2,6] the necessity applying logically – probability models.

Logical parts are interpreted using 9 operators – AND, OR, IF, $WHEN$, $THEN$, YES, NO, $START$, END. Each logic element of the danger is represented by single event, with analytical designation, response to its specification.

The probability part is presented by the indicator of the risks – the probability P for a given event appearance.

The combination from formalized events and logical operators interpreted descriptive model of a situation. On each event specific meanings are given. The envelopment and limitations of the analysis and assessment of the dangers is appropriated.

By appropriation of a single meaning of the events is defined mono situational danger's model, and when the values of the of the risks are represented the model is risk's mono situational. By interpretation with the all possible values of the events and their risks is defined poli-situational model of trans-border dangers.

Then formalized logical – probability risk's models for trans-border danger has the form:

$$start \rightarrow when \rightarrow R_{hazard\ event} \rightarrow then \rightarrow R_{hazard\ process} \rightarrow and \rightarrow R_{subjiecive\ source} \rightarrow or/and \rightarrow$$

$$R_{objiective\ source} \rightarrow then \rightarrow R_{prophy\ laxis} \rightarrow and \rightarrow R_{open/hazard\ emissuion} \rightarrow or \rightarrow R_{potential/hazard\ emission} \rightarrow and \rightarrow th$$

$$en \rightarrow R_{middle\ diffusion*1} \rightarrow and \rightarrow R_{level}\ level \rightarrow and \rightarrow R_{hazard\ emission} \rightarrow and \rightarrow R_{excess} \rightarrow and \rightarrow$$

$$R_{protection\ action} \rightarrow and \rightarrow R_{hazard\ imission} \rightarrow then \rightarrow R_{middle\ diffusion*2} \rightarrow and \rightarrow R_{impact\ space} \rightarrow and \rightarrow$$

$$R_{transborder\ diffusion} \rightarrow and \rightarrow R_{transborder\ level} \rightarrow and \rightarrow R_{transborder\ duration} \rightarrow and \rightarrow R_{impact\ obkjiect} \rightarrow and \rightarrow$$

$$R_{objiect\ sesnsitivity} \rightarrow and/or \rightarrow R_{reduction\ protection} \rightarrow and/or \rightarrow R_{personal\ protection} \rightarrow and \rightarrow R_{harm\ aspect} \rightarrow and \rightarrow$$

$$R_{harm\ location} \rightarrow and \rightarrow R_{harm\ weigth} \rightarrow and \rightarrow R_{harm\ dimension} \rightarrow and \rightarrow R_{harm\ recuperation} \rightarrow and \rightarrow$$

$$R_{harm\ compensate} \rightarrow then \rightarrow end. \tag{1}$$

The structure of the poli-situational model is shown at fig. 2. Each risk element is a product of all of its values. This is the way to formalize all possible situations.

The criticalities are sum of all risks. Five criticalities categories are introduced:

I. Factor criticalities $CR\langle factor \rangle$. They describe the sum of risk for partial danger factors appearance at the origin. In a training stage they are inserted as $CR\langle factor \rangle = \{(R_{Factor*1}),(R_{Factor*2}),...,(R_{Factor*k})\}$, whereas $R_{Factor*1}$ is a risk of appearance, for instance noise, $R_{Factor*2}$ - risk of appearance of nitrogen oxides and etc.

II. Indicator criticalities $CR\langle indicator \rangle$. Shows the product of all the risks for base indicators of the danger events in trans-border zones, appearance: $CR\langle indicator \rangle = \{(R_{Indikator*1}),(R_{Indicator*2}),...,(R_{Indicator*l})\}$, whereas $R_{Indicator*1}$ is the risk of natural or anthropogenic danger events appearance, $R_{Indicator*2}$ - the risk of appearance of a danger processes, $R_{Indicator*3}$ - risk for appearance of danger sources and actions and etc.

III. Component criticalities $CR\langle component \rangle$. They are indicator for the risks for divided appearance of danger events, actions and effects of the trans-border dangers: $CR\langle component * phenomenon \rangle = \{(R_{Pheno})\}$,

$CR\langle component * impact \rangle = \{(R_{Impact})\}, CR\langle component * effect \rangle = \{(R_{Effect})\}$,

където $R_{Pheno}, R_{Impact}, R_{Effect}$ are the risks for danger events, actions and effects.

IV. Differential criticalities $CR\langle differential \rangle$. They are indicator of the arrange of all the risks for combine appearance of: 1) danger phenomenons and effects - $R_{\frac{Pheno}{Impact}}$, 2) danger impacts and effects -

$R_{\frac{Impact}{Effect}}$, 3) danger phenomenons and effects - $R_{\frac{Pheno}{Effect}}$, whereas:

$$CR\langle differential \rangle = \left\{ (R_{\frac{Pheno}{Impact}}),(R_{\frac{Impact}{Effect}}),(R_{\frac{Pheno}{Effect}}) \right\}.$$

V. Integral criticalities $CR\langle integral \rangle$. They unify the risks of simultaneous appearance of danger phenomenons, impacts and effects: $CR\langle integral \rangle = \{(R_{Pheno}),(R_{Impact}),(R_{Effect})\}$. Whereas $R_{Pheno}, R_{Impact}, R_{Effect}$ are the risks of simultaneous appearance of danger phenomenons, impacts and effects.

On the base of the resumed dependencies must be determine the numerical values of each the criticalities $CR\langle factor\rangle$, $CR\langle indicator\rangle$, $CR\langle component\rangle$, $CR\langle differential\rangle$, $CR\langle int\,egral\rangle$.

Our experience [1,2,3,4,6,7] showed the most appropriate is using the vector method.
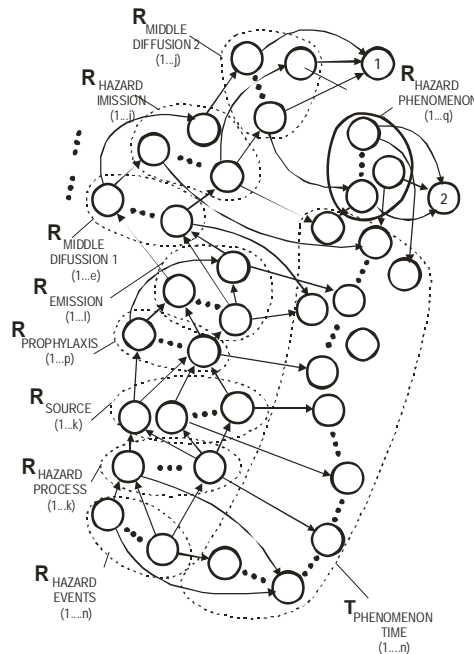
Fig. 2. Poli-situational model of risk for transborder danger phenomenon

It gives the possibility to take into account the relative part of each component. As a result we achieve a numerical value, reflecting all the composites. By contrast with insertion of temporal risks in other our investigations, in the presented one we use the real time T. This is the way to achieve objective vision about the dynamics of the changes of the criticality quantities.

The introduced above logical operators we use only for accurate creation of structural models of the trans-border dangers. When we apply the corresponding numerical dependencies, the results are not real. The composite products decrease the values of the resulted risks and it is not further possible to make an objective assessment.

The vector method is applied in a plane or spatial variant. In the plane situation each probability, respective risk, is according to a partial moment of appearance. It is introduced graphically as a point and its corresponding vector – fig. 3. This graph introduces the criticalities components $CR\langle component\rangle$.

It must be find the center of the points area of the danger phenomenons *Phenomenon area* (fig.3), danger impacts *Impact area* and danger effects *Effect area*. The vector of the three area centers is $\overset{\mathbf{\cdot}}{R}_{Pheno}$, $\overset{\mathbf{\cdot}}{R}_{Im\,pact}$ and $\overset{\mathbf{\cdot}}{R}_{Effect}$. The criticalities equals to the vectors and are

described with: $CR\langle component * phenomenon\rangle = \overset{\tiny\textbf{\textbar}}{R}_{Pheno}$,
$CR\langle component * impact\rangle = \overset{\tiny\textbf{\textbar}}{R}_{Im\,pact}$, $CR\langle component * effect\rangle = \overset{\tiny\textbf{\textbar}}{R}_{Effect}$.

Determination of the factor criticalities $CR\langle factor\rangle$, indicator criticalities $CR\langle indicator\rangle$, differential criticalities $CR\langle differential\rangle$ and integral criticalities $CR\langle int\,egral\rangle$ is made on the same method.

The component criticalities $CR\langle phenomenon\rangle$, $CR\langle impact\rangle$, $CR\langle effect\rangle$, according to above equations:

$$CR\langle phenomenon\rangle = \overset{\tiny\textbf{\textbar}}{R}_{Phenomenon} = \overset{\tiny\textbf{\textbar}}{R}_{\substack{Hazard\\event}} + \overset{\tiny\textbf{\textbar}}{R}_{\substack{Hazard\\Pr\,ocess}} + \overset{\tiny\textbf{\textbar}}{R}_{Source} + \overset{\tiny\textbf{\textbar}}{R}_{Pr\,ophrlaxis} +$$

$$\overset{\textbf{v}}{R}_{\substack{open/hazard\\emission}} + R_{\substack{potencial/hazard\\emission}} + \overset{\textbf{r}}{R}_{\substack{middle\\diffusion*1}} + \overset{\textbf{r}}{R}_{\substack{Hazard\\emossion}} + \overset{\textbf{r}}{R}_{Level} + \overset{\textbf{r}}{R}_{Excess} + \overset{\textbf{r}}{R}_{\substack{Pr\,otection\\action}},$$

(2)

$$CR\langle impact\rangle = \overset{\tiny\textbf{\textbar}}{R}_{Im\,pact} = \overset{\tiny\textbf{\textbar}}{R}_{\substack{Im\,pact\\middle}} + \overset{\tiny\textbf{\textbar}}{R}_{\substack{Im\,pact\\space}} + \overset{\tiny\textbf{\textbar}}{R}_{\substack{Im\,pact\\time}} + \overset{\tiny\textbf{\textbar}}{R}_{\substack{Im\,pact\\Level}} + \overset{\tiny\textbf{\textbar}}{R}_{\substack{Im\,pact\\object}} +$$

$$+ \overset{\textbf{r}}{R}_{\substack{Fljiect\\sensitivity}} + \overset{\textbf{r}}{R}_{\substack{Re\,duction\\protection}} + \overset{\textbf{r}}{R}_{\substack{Personal\\protection}},$$

(3)

$$CR\langle effect\rangle = \overset{\tiny\textbf{\textbar}}{R}_{Effect} = \overset{\tiny\textbf{\textbar}}{R}_{\substack{Harm\\aspect}} + \overset{\tiny\textbf{\textbar}}{R}_{\substack{Harm\\location}} + \overset{\tiny\textbf{\textbar}}{R}_{\substack{Harm\\weight}} + \overset{\tiny\textbf{\textbar}}{R}_{\substack{Harm\\dim\,ension}} +$$

$$\overset{\textbf{r}}{R}_{\substack{Harm\\recuperation}} + \overset{\textbf{r}}{R}_{\substack{Harm\\compensate}}.$$ (4)

The relations of these basic vectors to the time it is settled the impact ratios. KV_{Pheno}, $KV_{Im\,pact}$, KV_{Effect}. They introduce the degree of impact of the probability or the time of the criticalities to appear. The spatial variant of the component criticalities is proposed at fig. 4. The relationships of $\overset{\tiny\textbf{\textbar}}{R}_{Pheno}$, $\overset{\tiny\textbf{\textbar}}{R}_{Im\,pact}$, $\overset{\tiny\textbf{\textbar}}{R}_{Effect}$ to $\overset{\tiny\textbf{\textbar}}{R}_{Integral}$ are inserted as importance ratios KZ_{Pheno}, $KZ_{Im\,pact}$, KZ_{Effect}. They
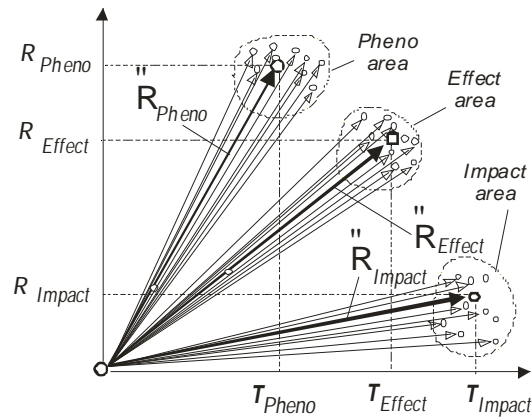
Fig. 3. Plane risks presentation

show the importance degree of the danger phenomenon, impacts and effects upon the situational criticality. When the monitoring time is put - T and risks - R the following regions of variation are gotten $3,3166 \le R_{Pheno} \le 15,5563$, $8,0000 \le R_{Impact} \le 11,3137$, $6,0000 \le R_{Effect} \le 8,4852$, $14,8660 \le R_{Integ} \le 21,0236$.
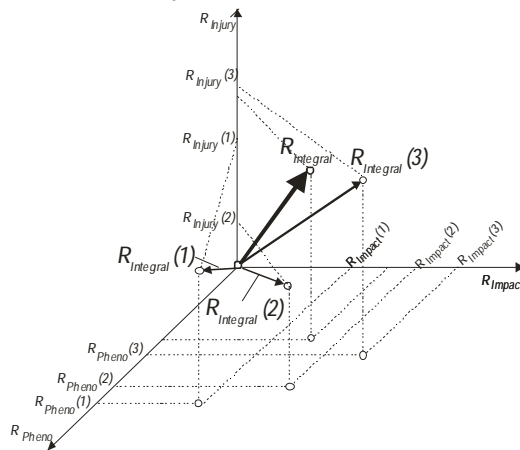
Fig.4. Spatial interpretation of the integral risk $R_{Integ} = CR\langle integral\rangle$ from 3 risk situations

$$R_{Integ}(1),$$
$$R_{Integ}(2), R_{Integ}(3)$$

The maximum values are achieved when the biggest values of the probabilities are attached, being 1, and 1 for monitoring of elements appearance, for instance 1 month, 1 year and etc.

Each zone is divided to 5 equal parts, corresponding to 5 degrees of the criticality – very little, little, middle, large and very large. These degrees allow filtering the criticalities. This leads to easy navigation and juxtaposition of the values of the criticalities from different situations.

From the shown above the following conclusions could be made:

1. The models of the dangers, risks and criticalities in the environment of the trans-border zones were introduced. Applying it it is possible to set the numerical values, assessments and juxtapositions.

2. An appropriate method was proposed for the risk of border dangers appearance modeling, revealing the essence of the certainty of different activities in the border areas.

3. A combine logically and probability modeling, that result possibilities for reproducing of variety of border's danger situations, was implemented.

4. Applying the method for vector numerical investigation it was achieved mathematical model, used for continuous determination of the risk and criticalities. The spheres of alternation of the risks were settled and 5 criticalities degrees were introduced, able to apply for fast juxtaposition.

References

1. Владимиров, Л. Екологични трансгранични критичности – структура и специфика. Свищов, Стопанска академия "Д. А. Ценов", Юбилейна международна научно-практическа конференция "Хоризонт пред икономическото знание и бизнеса", 7–8.10.2010 г.

2. Владимиров, Л. Рискметрия в екологичната сигурност. Монография. Варна, Варненски свободен университет "Черноризец Храбър", 2009, с. 279.

3. Владимиров, Л. Количествена оценка на критичността. София, Българска академия на науките, Втора научна-практическа конференция по управление в извънредни ситуации. 9.11. 2007, с. 172–179.

4. Владимиров, Л., В. Томов. Каузални риск модели. Част II. Таксономичен модел. Русе, Русенски университет "Ангел Кънчев", Сборник Научни трудове' 2007, 90–95.

5. Владимиров, Л. Екологични трансгранични критичности – структура и специфика. Свищов, Академично издателство "Ценов", Стопанска академия "Д. А. Ценов", Сборник "Хоризонт 2020 пред икономическото знание и бизнеса", 2010, том 1, с. 119-124.

6. Владимиров, Л. Аналитичен модел на екологичните трансгранични критичности. Свищов, Акад. изд. "Ценов", Стопанска академия "Д. А. Ценов", Сборник "Хоризонт 2020 пред икономическото знание и бизнеса", 2010, том 1, с. 112–118.

7. Vladimirov, L. Criticality Assessment. Ecologica International, №50, XIV, 2007, p. 23–29;

Authors:
Assoc. prof. PhD Lyubomir Vladimirov Vladimirov
Chief Ass. Nikolai Iordanov Kovachev
Chief Ass. Plamen Mihailov Manev
Chief Ass. PhD Ventsislav Dobrinov
Rousse 7017, Str. „Studetnska" № 8, University of Ruse „Angel Kanchev,
E-mail: lvvladimirov@uni-ruse.bg, nkovachev@uni-ruse.bg, pmanev@uni-ruse.bg, vdobrinov@uni-ruse.bg

INFORMATICS IN ENGINEERING AND TECHNOLOGIES

COMPARISON OF DECISIONS FOR EMULATION DEVICES TO INDICATING AND MONITORING IN MANEUVERABLE SIMULATOR

Stefan Borisov

Abstract: *The article presents specifics of the construction of instruments - software and hardware designed to use on the bridge in full mission ship simulators. Discussed are the possible software solutions for emulating devices. The advantages and disadvantages of these solutions, possibility to establish communications with other platforms in a network environment and some results of practical realization of these decisions are discussed. The results of implementation of selected software method to do this emulation on touch screen monitors are analyzed.*

Key words, *emulation devices, HMI development tools, command bridge, physical modeled equipment, Network-Published Shared Variable, shared variable engine, OPC interface.*

Part of the problem of creating a simulator is modeling the systems for management and monitoring of the bridge of the simulated object. Part of this modeling is designing a model of the ship command bridge and in particular emulation devices for indication and control. In connection with the design of simulator at Bulgarian Ship Hydrodynamics Centre(BSHC) meetings were held for project concept discussion.Based on discussions during meetings with specialists and partners, it was decided to reproduce almost all physical equipment of the real bridge left and right steering stands of the most used pushers from Bulgarian River Shipping J.S.Co [5].
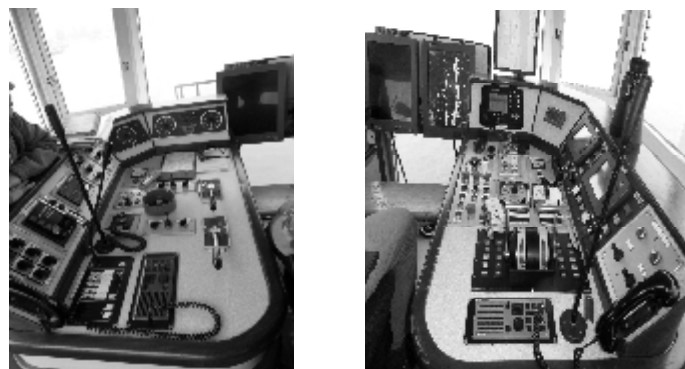
Photo 1.1. Steering stand (left and right) of the real pusher

Photo 1.1 shows the right and left control console of the real pusher bridge. On the photo 1.2 are shown the corresponding physical models. The models of the bridge control stands are also a little bit larger than the real one to enable the mounting of the touch screen monitors on which are virtually modeled some of the real equipment on the control consoles.

Photo 1.2. The right and left steering stands of the simulator

In the construction of the model of the pusher command bridge was decided to use two approaches for its emulation: software and hardware. On the hardware approach – the equipment on the bridge is simulated physically. Photo 1.3 shows one example - Trust Levers for both engines propeller revolutions setting. The original Trust Levers are shown on the photo 1.3 left.
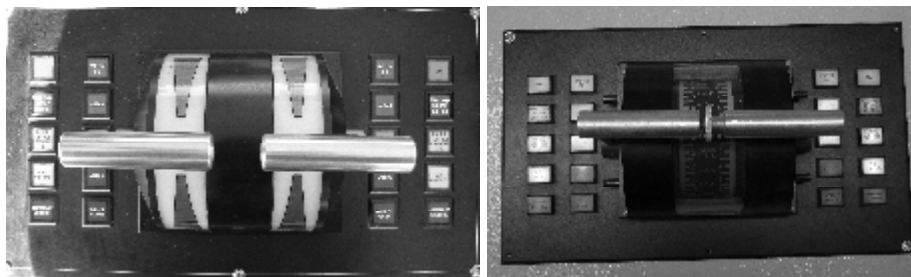
Photo 1.3. The machine telegraph models (Real-left; Model – right)

For software modeled instruments are used touch screen monitors.The basic concept of object-oriented development is for a computer program to be created from a collection of objects where each

element is capable of receiving messages, processing data and sending messages and data to other objects. In a graphical HMI development tool, objects are the various components that are used to create a display. These elements vary greatly in complexity, ranging from a simple line or image to more elaborate elements such as an altimeter in an aircraft or a speedometer. Other types of objects such as data calculators used to manipulate data; data objects that can be used as variables for information storage and device input objects which are able to capture input from the user do not have a visual representation, but also contribute to the functionality of a display. Next-Generation HMI development tools offer all of these basic elements but also offer the possibility to create new objects to fulfill more advanced requirements. These new components can be created graphically within the user-interface. Once they are created, they reside within the user interface like all other elements that were originally delivered with the software. They also share the same type of interface namely a set of properties to be able to configure them, operations that can be called to send messages to the object and events that the object can generate to communicate with the rest of the application [3]. On Photo 1.4 is showing an example of an object - construction of a flank rudders angle indication system, modeled on 12" touch screen LCD monitor. Such an object would have properties to indicate its position on the screen, the maximum angle of the instrument and the current angle position of the flank rudders. Once its interface is defined, its internal definition is created using a knob and needle to indicate the angle position, text and lines to create a scale behind the needle.With all of these elements in place, control logic can finally be added to send the incoming data to the needle.
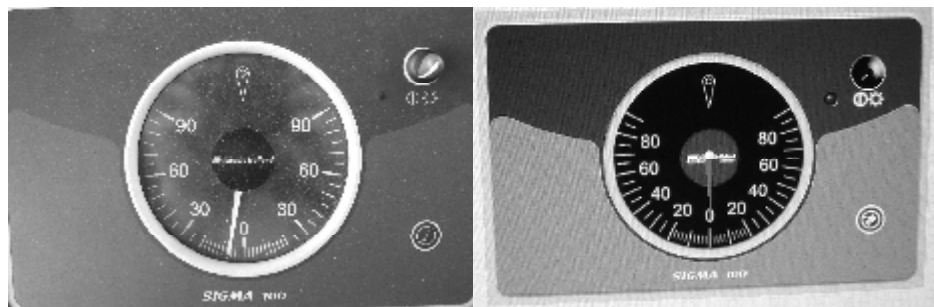
Photo 1.4. Indicators for flank rudders angle (Real – left; Model – right emulated with 12" touch screen monitor)

On Photo 1.5 is show another example of software modeled instruments - Communication Intercom and UHF radio modules. These modules are simulated in LabView programming environment on 17" touch screen LCD monitor.

Photo 1.5 Intercom and UHF radio (Real-left; Model – right simulated on 17" touch screen LCD monitor)

For emulation of devices BSHC had the possibility to choose between two alternatives- VAPS and LabView. VAPS XT is a product of Presagis company. VAPS XT is a full-featured and advanced software tool for the rapid development of dynamic, interactive graphical interfaces for safety and mission-critical avionics and embedded displays. VAPS XT offers unparalleled control and flexibility in the design of dynamic, interactive, real time human machine interfaces (HMIs). Ideal for the design of all types of embedded display graphics ranging from digital and analog instrumentation to next-generation menu-based displays. The product has been created by HMI tool experts with years of industry expertise and has been developed through close consultation with tier-1 embedded manufacturers [1].

Main advantages of this product
• Object-Oriented Development Environment
• Encapsulation of legacy object code and communication Interfaces

VAPS XT is an object-oriented model-based development tool that enables users to create the visual appearance, logic, and data connectivity of their displays in a user-friendly environment.
• Graphical logic editing facilities

144

Defining application logic or the control logic in an object definition can be done at many different levels. A first layer of logic is to describe the flow of data between the components that make up the display or the object. This type of connection is simple to make but also has limitations since they are always active and cannot easily react if certain thresholds have been exceeded. The solution is to use more complex logic in the form of a state chart diagram. They are used to visually represent finite state machines, which are behavior models composed of states, transitions and actions.

Main disadvantages

• Designed primarily to create dynamic, interactive graphical interfaces for safety and mission-critical avionics and embedded displays.

• No possibility for data acquisition of the signals and events from the physical modeled equipment.

VAPS XT can communicate with other application, running on the same machine, and across a network. VAPS XT Communications interface which lets you build applications to communicate with:

• several VAPS XT applications

• external applications running on the same host

• external applications running on other hosts (using Ethernet)

But the package which is bought by BSHC is not including this communication interface.

The second alternative was the software package LabView from National Instruments. LabView, short for Laboratory Virtual Instrument Engineering Workbench, is a programming environment in which user create programs using a graphical notation (connecting functional nodes via wires through which data flows); in this regard, it differs from traditional programming languages like C, C++, or Java, in which user program with text. LabView offers more flexibility than standard laboratory instruments because it is software-based. The users, not the instrument manufacturer, define instrument functionality. The computer, plug-in hardware - data acquisition (DAQ) devices and LabView comprise a completely configurable virtual instrument to accomplish the tasks [4]. Thanks to this feature can be done data acquisition of the signals coming from physical modeled equipment. In this case Trust Levers output voltage, proportional to the engine's revolution, is taken out from potentiometers, mounted on the shaft. The outputs are connected to the analog inputs (AI) of NI USB DAQ Card. On the Trust Levers panel there are buttons for emergency alarm and control of the engines same as on the real pusher. The buttons that switch the engine speed control between telegraph and thrust lever are placed on this panel too. The

events generate from this buttons can be processed as the outputs from this buttons are connected to the digital inputs (DI) of NI USB DAQ Card.

Main advantages

• Object-Oriented Development Environment

In LabView are used custom controls and indicators to extend the available set of front panel objects. It is possible to create custom user interface components for an application that vary cosmetically from built-in LabView controls and indicators. For example, it is possible to change the size, color, and relative position of the elements of a control or indicator and import images into the control or indicator.

• Graphical logic editing facilities

Simplify complex state-based applications with the NI LabView Statechart Module. This LabView add-on provides a high level of abstraction for designing applications using states, transitions, and events.

NI LabView can communicate with other application, running on the same machine, and across a network. NI LabView Communications interface which lets you build applications to communicate with:

• several LabView programs within the same application instance
• external applications running on the same host
• external applications running on other hosts (using Ethernet)

LabView provides access to a wide variety of technologies for creating distributed applications. The shared variable is a major step forward in simplifying the programming necessary for such applications. First option can be done using global variables to access and pass data among several LabView programs. This option can be done also using Single-Process Shared Variable. The single-process variables is use to transfer data between two different locations on the same program that cannot be connected by wires, such as parallel loops on the same program, or two different programs within the same application instance. The underlying implementation of the single-process shared variable is similar to that of the LabView global variable. The main advantage of single-process shared variables over traditional global variables is the ability to convert a single-process shared variable into a network-published shared variable that any node on a network can access. Second and third option can be done with Network-Published Shared Variable. Using the network-published shared variable, user can write to and read from shared variables across an Ethernet network. The networking implementation is handled entirely by the network-published variable. In addition to making the data available over the network, the network-published shared variable adds many features not available with the single-process shared variable. To provide this additional

146

functionality, the internal implementation of the network-published shared variable is considerably more complex than that of the single-process shared variable. The NI Publish and Subscribe Protocol (NI-PSP) is a networking protocol optimized to be the transport for Network Shared Variables. The shared variable engine (SVE) is a software framework that enables a networked-published shared variable to send values through the network. On Windows, LabView configures the SVE as a service and launches the SVE at system startup. In order to use network-published shared variables, an SVE must be running on at least one of the nodes in the distributed system. Any node on the network can read or write to shared variables that the SVE publishes [2]. The SVE functions as an Object Linking and Embedding for Process Control (OPC) 2.x and OPC 3.0 compatible server. LabView uses the National Instruments Variable Engine OPC interface to display all numeric, boolean, and string shared variables that the user deploy to the SVE. Figure 1 shows the memory that the SVE uses after LabView deploys 500 and 1000 shared variables of the specified types to it. The graph shows that the type of variable does not affect the memory usage of the deployed shared variables significantly. Note that these are nonbuffered variables.
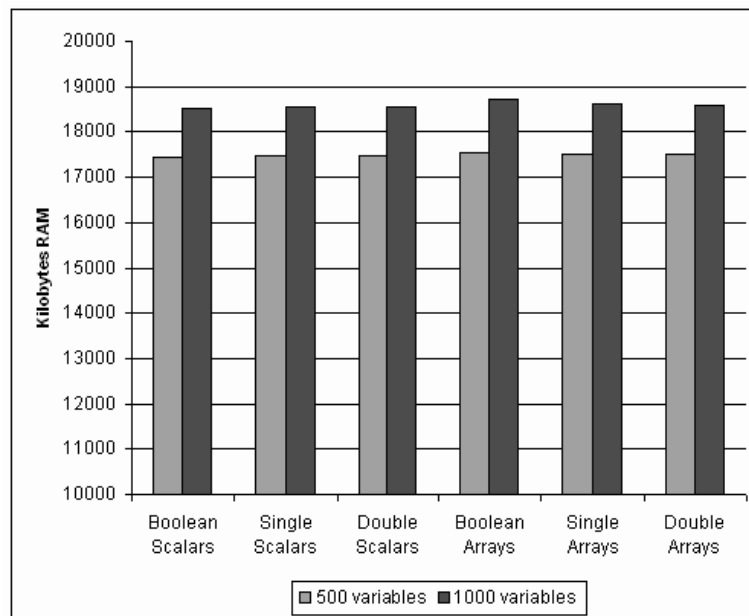
Figure 1. Memory Usage of Network-Published Shared Variables with Different Data Types

In our project we use less than 500 variables which mean that memory usage of network-published shared variables was significantly reduced. The SVE updates the shared variables when the server sends new data to the SVE. Servers also update each output when the Human Machine Interface (HMI) application writes that shared variable value and they handle and report communications and device errors. It is possible to use any OPC 2.x or OPC 3.0 client to view and modify OPC data that the OPC server publishes. To create connection with other node on the network for example C++ application is use OPC interface Beharrell - OPC DA Client SDK. OPC DA Client SDK is a complete toolkit for standard OPC Data Access 2.x, providing wrapper classes to hide much of the complexity of OPC.

CONCLUSION

After analyzing the possibilities of the products LabView and VAPS we find that for BSHC simulator LabView is the best solution due to the following features:

• Easy modeling – Huge amount of hardware and software tools for creating control instruments and indicators on the bridge of the full mission ship simulator.

• There is a wide range of modules with analog and digital inputs and outputs for signals and data acquisition from physical modeled instruments on the bridge, which can be managed with LabView.

• LabView has a rich set of built-in functions for capturing and exchange of events and data from software modeled equipment.

• Built-in integration of OPC in LabView. This interface allows data exchange with other software packages used to generate and manage virtual reality in simulator.

• In LabView there is a lot of virtual instruments (VI) for data exchange and integration with other devices using standard protocols - NMEA, USB, RS, TCP / IP and others.

• In program created with LabView can be easy incorporation programs developed with other programming environments - Matlab, Visual Studio e.t.c.

Literature
1. www.presagis.com/products_services/products/embedded-graphics/hmi_modeling_and_display_graphics/vaps_xt/
2. www.ni.com/white-paper/4679/en
3. THE MODERNIZATION OF HMI DEVELOPMENT TOOLS TO SUPPORT THE ENTIRE DEVELOPMENT PROCESS by Yannick Lefebvre

4 LabVIEW for Everyone: Graphical Programming Made Easy and Fun, Third Edition By Jeffrey Travis, Jim Kring
5. „Design of Inland Navigation Full Mission Bridge Simulator", финален отчетен доклад по проект BUL/017/07 , Център по хидро- и аеродинамика към ИМСТ „А.Балевски"-БАН, юли 2011

Author:
senior Expert eng. **Stefan Hristov Borisov**, Ph.D.
9003 Varna Street, "William Froude" № 1
Asparuhovo, PO Box 58
Bulgarian ship hydrodynamics centre to Institute of metal science, equipment and Technologies "Acad. A.Balevski "-BAS
Tel: 052 370501\444, e-mail:s.borisov@bshc.bg

DIGITAL RESOURCES FOR LANGUAGE TECHNOLOGIES AND THEIR APPLICATIONS

Ludmila Dimitrova

Abstract: The paper describes applications of information technologies for development of special kinds of digital resources, namely, digital multilingual corpora. Digital corpora with Bulgarian (Bulgarian-Polish and Bulgarian-Polish-Lithuanian) and their applications as resources for language technologies are presented. Some applications in machine translation, contrastive studies and education are also described.

1. Introduction

Multilingual corpora are large repositories of natural language data with an important role in natural language processing (NLP). These digital resources are a valuable multilingual dataset for language engineering research and development for the Bulgarian language.

The paper briefly reviews applications of the information and knowledge technologies for development of digital multilingual corpora which are a special kind of digital resources. This paper describes two digital corpora: Bulgarian-Polish and Bulgarian-Polish-Lithuanian. These corpora are currently under development, they are constantly growing with new features being added. The Bulgarian-Polish aligned corpus will be soon represented via Internet as a web-based tool with a wide spectrum of features for practical applications. The corpus will be freely available for research on the web with a multilingual (in Bulgarian, Polish, and also in English) interface.

Every web-based corpus is a resource combining a number of features that together make it unique and useful tool not only for language studies, but also for researchers of many fields. These features include: rich linguistic content (parallel texts in two Slavic languages and a Baltic one, for example, in this case), annotation (mark-up at two levels: paragraph and sentence), search query (advanced possibilities for combining many search criteria), display of the search results in an intuitive and simple interface, advanced results handling (concordances, collocations, etc.).

Both corpora (see [2], [4]) are under development. They are the result of close collaboration between the partners in the joint research project "Semantics and Contrastive Linguistics with a Focus on a Bilingual Electronic Dictionary" between IMI-BAS and ISS-PAS, coordinated by L. Dimitrova and V. Koseska.

Reasons for the development of the two corpora

The Bulgarian-Polish and Bulgarian-Polish-Lithuanian corpora are developed due to the need for research material for contrastive studies in these three languages.

Two of the languages, Bulgarian and Polish belong to the Slavic language family: Bulgarian belongs to the South-Slavic language group and Polish – to the West-Slavic language group, Lithuanian belongs to the Baltic language group. Linguistic and contrastive studies of these languages can be carried out based on bilingual digital resources (corpora and dictionaries). Furthermore, the first Bulgarian-Polish parallel corpus serves as a main source of vocabulary for the Bulgarian-Polish digital dictionary.

Corpus annotation

Corpus annotation is the process of adding linguistic or structural information to a text corpus ([7], [9]). One common type of annotation is the addition of labels or tags that indicate the word class for the words in the text. This is the so called part-of-speech tagging (or POS tagging).

Apart from POS tagging, there are other types of annotation, for example, structural annotation, which corresponds to different structural levels of a corpus or text. Written texts contain a number of different structural forms or divisions. Novels have a complex hierarchy and are divided into parts and chapters, newspapers are divided into sections, reference works – into articles, etc. The most common division in this hierarchy is the paragraph.

Some texts in the ongoing version of the Bulgarian-Polish corpus are annotated at paragraph level, others are aligned and therefore annotated at the segment level (usually sentence level). The standard markers <p> and </p> for the paragraph's boundaries, <seg> and </seg> for the segment's boundaries, are employed. This annotation allows texts in both languages (Bulgarian/Polish and vice versa) to be aligned at paragraph (<p> level) or at segment level in order to produce aligned bilingual corpora. The <p> level alignment allows the drawing of a broader context in the languages. In other words, there is the opportunity – thanks to the broader context – to study more precisely the meanings of word-forms in each language.

2. MULTEXT-East corpus – model for corpus development

The formal model that is used in the development of the first Bulgarian-Polish corpus is the MULTEXT-East corpus.

For Bulgarian language the first parallel corpus has been produced as a part of the MULTEXT-East (MTE) corpus[1] consisting of George Orwell's "1984" in English and its translations in six MTE languages. The alignment of sentences from the English original with the Bulgarian text is also a part of the MTE corpus ([1]).

MTE project is building an annotated multilingual corpus, composed of three major parts:

• Parallel Corpus,

• Comparable Corpus,

• Speech Corpus (a small one) of spoken texts in each of the six languages, comprising forty short passages of five thematically connected sentences, each spoken by several native speakers, with phonemic and orthographic transcriptions.

2.1. Parallel MTE corpus

A parallel text is a text placed alongside its translation or translations. Large collections of bi- or multilingual parallel texts are parallel corpora.

Multilingual parallel corpus, based on George Orwell's novel "1984" in the English original and the six translations in Bulgarian, Czech, Estonian, Hungarian, Romanian and Slovene of the novel, was developed. The parallel corpus is produced as a well-structured, lemmatized, CES-corpus [1].

2.2. Annotated MTE parallel corpus

The texts were automatically annotated for tokenization, sentence boundaries, and part of speech (POS) annotation (with morpho-syntactic descriptors MSD), using the project tools, and validated for sentence boundaries and alignment. Bulgarian text of the MTE Orwell's "1984" corpus contains word-level morpho-syntactic mark-up (undisambiguated lexical information for 156002 words, 156002 occurrences of MSD, and disambiguated lexical information for the 86020 words of the novel). The following figure shows disambiguated lexical information for the Bulgarian words "*ден*" and "*бе*" from the first sentence of the novel: *Априлският ден бе ясен и студен, часовниците биеха тринайсет часа.* (*It **was** a bright cold **day** in April, and the clocks were striking thirteen.*)

```
…………..
<tok type=WORD from='Obg.1.1.1.1\12'>
<orth>ден</orth>
<disamb><base>ден</base><ctag>NCMS-N</ctag></disamb>
<lex><base>ден</base><msd>Ncms-n</msd><ctag>NCMS-N</ctag></lex>
</tok>
<tok type=WORD from='Obg.1.1.1.1\16'>
<orth>бе</orth>
<disamb><base>съм</base><ctag>VAIA3S</ctag></disamb>
<lex><base>бе</base><msd>Qgs</msd><ctag>QG</ctag></lex>
<lex><base>съм</base><msd>Vaia2s</msd><ctag>VAIA2S</ctag></lex>
<lex><base>съм</base><msd>Vaia3s</msd><ctag>VAIA3S</ctag></lex>
</tok>
………….
```

Fig. 1. An excerpt from the annotated Bulgarian MTE parallel corpus

2.3. Aligned MTE corpus

An aligned corpus is a multilingual (at least bilingual) parallel corpus. It is a result of the process of parallel text alignment that aims to produce a set of corresponding sentences (original and its translation(s)) in both or more parts of the parallel text (one of the most well-known example of parallel text alignment is inscribed on the famous Rosetta Stone). The result of the alignment of two parallel texts is a merged document, called bi-text, composed of both source- and target-language versions of a given text that retains the original sentence order. The alignment between the English original of the Orwell's "1984" and translations in each of the six MTE languages produces six pair-wise alignments comprising the MTE aligned corpus. Several different software tools were used for producing such corpus. For Bulgarian language the first parallel aligned corpus has been produced as a part of the MTE aligned corpus by means of the Vanilla Aligner software.

The next table shows excerpts of the Bulgarian-English aligned "1984" texts:

Table 1. Bulgarian-English aligned sentences

1-1 aligned sentences	<Obg.1.1.3.4>Звукът от апарата (наричаше се телекран) можеше да бъде намален, но нямаше начин да се изключи напълно. <Oen.1.1.3.4>The instrument (the telescreen, it was called) could be dimmed, but there was no way of shutting it off completely.
1-2 aligned sentences	<Obg.1.1.24.8>Изпитваше дълбок интерес към него не само защото беше заинтригуван от контраста на изисканите маниери с телосложението му на борец, а много повече заради стаената увереност -- или навярно не толкова увереност, колкото надежда, -- че политическата правоверност на О'Брайън не е изрядна. <Oen.1.1.25.8>He felt deeply drawn to him, and not solely because he was intrigued by the contrast between O'Brien's urbane manner and his prize-fighter's physique.<Oen.1.1.25.9>Much more it was because of a secretly held belief -- or perhaps not even a belief, merely a hope -- that O'Brien's political orthodoxy was not perfect.
2-1 aligned sentences	<Obg.1.1.39.1>Това ставаше винаги през нощта. <Obg.1.1.39.2> Арестуваха неизменно през нощта. <Oen.1.1.41.1>It was always at night -- the arrests invariably happened at night.

2.4. Comparable MTE corpus

For each of the six MTE Central and East-European languages, a comparable corpus was developed. It included two subsets of at least 100 000 words each, consisting of fiction, comprising a single novel or excerpts from several novels; and newspapers. The data was comparable across the six languages, only in terms of the number and size of texts. The entire MTE multilingual comparable corpus was prepared in CES format, manually or using ad-hoc tools.

The Bulgarian subset of MTE comparable corpus includes fiction (texts from contemporary Bulgarian literature, 97251 words), non-fiction and newspapers (newspaper excerpts, 96538 words) subsets. It was annotated manually at the paragraph level, tagged with sub-paragraph mark-up (abbreviations, dates, names, direct/indirect speech).

```
<p>
        <q rend="PRE=mdash">Какво е това, дето го чете момичето,
        госпожо, неясно ми се вижда, да не й повлияе зле.
        </q>
</p>
<p>А тя ми отвърна като знаеща:</p>
<p>
        <q rend="PRE=mdash"> Философия,
                <name type="person">Йо</name> .
        </q>
</p>
```

Fig. 2 An excerpt from "Bulgarian fiction", a part of the MTE comparable corpus

3. Description of the Bulgarian-Polish corpus

3.1. Linguistic contents and numbers
The corpus contains texts (data) in the national languages: Bulgarian and Polish. It is constantly growing, since new data are adding, or are planned for adding (other texts are in various stages of completion).

At the moment it contains digital data approximately of 5 million words. Eventually, the volume of the corpus will rise to around 7-8 million words.

All texts in the corpus are texts published in and distributed over the Internet or were given by the authors for research purposes only.

3.2. Bulgarian–Polish Parallel and Comparable Corpora
The Bulgarian–Polish corpus consists of two parts: a parallel and a comparable corpus. The parallel corpus contains literary texts and texts of documents in both languages, whereby the translation correspondence is one-to-one. The Bulgarian–Polish corpus is annotated according to the standards of Text Encoding Initiative (TEI) for such kind of language resources. The entire Bulgarian–Polish corpus will be prepared in CES format, manually or using ad-hoc tools, and will be annotated for sentence (<s>, </s>) and paragraph (<p>, </p>) boundaries (sentence **S**-level and paragraph **P**-level).

The **Bulgarian–Polish parallel corpus** includes two parallel sub-corpora:
1) core Bulgarian–Polish corpus
2) translated Bulgarian–Polish corpus.

The core Bulgarian–Polish corpus consists of

• original texts in Polish - some novels by Ryszard Kapuściński, some fiction novels by Stanisław Lem and other Polish writers and their translation in Bulgarian,

• original texts in Bulgarian - short stories by Bulgarian writers and their translation in Polish.

The translated Bulgarian–Polish corpus consists of texts in Bulgarian and in Polish of brochures of the EC, documents of the EU and the EU-Parliament, published in Internet; Bulgarian and Polish translations of Antoine de Saint-Exupery's "The Little Prince"; Bulgarian and Polish translations of Michael Bulgakov's "Master and Margarita".

A part of the parallel texts is annotated at paragraph level.

Table 2. Alignment at the paragraph level: Stefan Żeromski's *Ashes* (vol. 1, part 1)

Polish:	Bulgarian:
** **Psy ucięły. Zaraz potem drugi głos, bliższy Rafała, odpowiedział jednokrotnie tym samym sposobem.** **	** **Кучетата млъкнаха. Веднага след това друг глас, по-близко до Рафал, отговори еднократно по същия начин.** **
Polish:	Bulgarian:
** **Młody myśliwiec jeszcze przez czas pewien leżał na ziemi, pękając ze złości: Po chwili jednak zerwał się na równe nogi, strzepnął śnieg z siebie, odszukał w krzakach pojedynkę. Wytarł oczy i, na podobieństwo sarn skacząc przez choiny, pomknął na dół.** **	** **Младият ловец лежа още малко на земята, позеленял от яд. Но после изведнъж скочи на крака, изтупа снега от себе си и потърси пушката в храстите. Избърса очи и скачайки като сърна през младите елички, полетя надолу.** **

The **Bulgarian–Polish comparable corpus** includes texts in Bulgarian and Polish with the text sizes being comparable across the two languages: excerpts from textual documents, shown in internet, excerpts from several original fiction, novels or short stories. Some of the Bulgarian texts in the comparable Bulgarian–Polish corpus are annotated on **P**-level and **S**-level, according to the international standards ([8]). Bulgarian–Polish comparable corpus includes novels of Bulgarian authors, namely: Dimitar Talev, Dimitar Dimov, Pavel Vezhinov, Yordan Radichkov.

3.3. Bulgarian-Polish Aligned Corpus

For a parallel corpus to be useful, it must be treated with a special program for "alignment". The term "alignment" means the process of connecting pairs of words, phrases, terms or sentences in texts from different languages that are translated equivalents of each other. "Alignment" is a form of annotation carried over parallel corpora to facilitate construction and evaluation of translation models stored in memory and used in support of computer-assisted translation.

Although many parallel corpora are manually "aligned", automatic "alignment" forms the core of parallel corpora processing and tool development for "alignment" with a high degree of accuracy.

Two language-independent freely-available software tools were used to align texts of the Bulgarian–Polish parallel corpus: MT2007 Memory Translation computer aided tool (TextAlign), and Bitext Aligner/Converter (Bitext2tmx aligner). TextAlign is available at http://mt2007-cat.ru/index.html. This software package segments and aligns corresponding translated sentences, contained in two rich text

format files. Bitext2tmx aligner (available at http://bitext2tmx. sourceforge.net) is a Java application. It works on any Java supported operating system (e.g. Windows, Linux, Mac OS X, Solaris), and is released under the GNU General Public License. Bitext2tmx aligner is a software tool that segments and aligns corresponding translated sentences, contained in two plain text files. These software packages have applications in computer-assisted translation. Both tools align bilingual texts without bilingual dictionaries, but the human editing is obligatory. The resulting aligned texts are similar.

The aligned corpus includes texts of Polish novels: Stanisłav Lem's Solaris and Return from the Stars, Ryszard Kapuściński's The Shadow of the Sun, Another Day of Life, and The Soccer War; and Stefan Żeromski's Ashes and their Bulgarian translations.

Table 3. Alignment at the sentence level: R. Kapuściński's *The Soccer War*

1-1 aligned sentences	**<tu tuid="0000000106">**
	    <tuv xml:lang="Polish">
	      <seg>Kazali mi czekac.**</seg>**
	    </tuv>
	    <tuv xml:lang="Bulgarian">
	      <seg>Казаха ми, че трябва да чакам.**</seg>**
	    </tuv>
	</tu>

1-2 aligned sentences	**<tu tuid="0000000125">**
	    <tuv xml:lang="Polish">
	      <seg> Znalazłem sie w straszliwych ciemnosciach, w ciemnosciach gestych, zbitych, nieprzeniknionych, jakby czarna i gesta maz zalała mi oczy, nie widziałem dosłownie nic, nawet wyciagnietych przed soba rak.**</seg>**
	    </tuv>
	    <tuv xml:lang="Bulgarian">
	      <seg> Потънах в страшна тъмнина, в непроницаем, непрогледен мрак, сякаш някаква гъста чернилка бе заляла очите ми. Не виждах абсолютно нищо, даже протегнатите си ръце.</seg>
	    </tuv>
	</tu>

2-1 aligned sentences	**<tu tuid="0000000193">**
	    <tuv xml:lang="Polish">
	      <seg> Był wystraszony, bo choc jego ojciec pochodził z Hondurasu, matka była obywatelka Salwadoru. Jako mieszaniec znalazł sie w kregu podejrzanych.**</seg>**
	    </tuv>
	    <tuv xml:lang="Bulgarian">
	      <seg> Той беше изплашен, защото макар и баща му да бе от Хондурас, майка му беше гражданка на Салвадор и като мелез той бе в кръга на заподозрените.**</seg>**
	    </tuv>
	    </tu>

4. Bulgarian-Polish-Lithuanian Parallel, Aligned and Comparable Corpora

4.1. Trilingual Parallel and Aligned Corpora

The **trilingual parallel corpus** contains more than 1 million words up to now. A part of the parallel corpus comprises original literary texts (fiction, novels, and short stories) in one of the three languages with translations in the other two, and texts of brochures of the European Commission, official documents of the European Union and the European Parliament, available through the Internet. The remaining part of the parallel corpus comprises texts in other languages translated into Bulgarian, Polish, and Lithuanian.

Thus, the texts are classified as either

• Core – original literary texts (fiction, novels, and short stories) in one of the three languages with translations in the other two, aligned at the paragraph and sentence levels, for example refer to Table 4.

• Translations – texts in other languages translated into Bulgarian, Polish, and Lithuanian: novels translated from other languages, aligned at the paragraph level (Antoine de Saint-Exupéry's Le Petit Prince, Michael Bulgakov's Master and Margarita), specialized texts of brochures and official documents of the European Commission, European Union, available through the Internet.

The recent result of the collaborative work is the aligned Bulgarian-Polish-Lithuanian corpus. At the first stage of the aligned process the align software tool were used to align the original text, for example Stanislav Lem's Solaris in Polish, and its Bulgarian translation. At the second stage the procedure is repeated with the input pair being the original Polish text and its Lithuanian translation. At the third stage, after a comparison of the two output aligned texts, Polish-Bulgarian and Polish-Lithuanian, the alignment end up with a sequence of triples: a sentence in Polish and its translations in Bulgarian and Lithuanian.

The following example presents an excerpt from the aligned at the sentence level texts of Stanislav Lem's Solaris (using TextAlign):

Table 4. Alignment at the sentence level

| 1-1-1 alignment sentences | `<tu tuid="0000000011">`
    `<tuv xml:lang="polish">`
      `<seg>`Widziałem już seledynowy kontur jedynego wskaźnika.`</seg>`
    `</tuv>`
    `<tuv xml:lang="bulgarian">` |

	`<seg>`Вече различавах светлозелените контури на универсалния указател.`</seg>`
	`</tuv>`
	`<tuv xml:lang="lithuanian">`
	`<seg>`Jau išskyriau žalsvus universalaus indikatoriaus kontūrus.`</seg>`
	`</tuv>`
	`</tu>`

1-1-2 alignment sentences	`<tu tuid="0000000024">`
	`<tuv xml:lang="Polish">`
	`<seg>`Niebo tych stron Galaktyki nic mi nie mówiło, nie znałem ani jednej konstelacji, w wąskim okienku trwał roziskrzony kurz.`</seg>`
	`</tuv>`
	`<tuv xml:lang="bulgarian">`
	`<seg>`Небето по тия места на Галактиката не ми говореше нищо, не познавах нито едно съзвездие.`</seg>`
	`</tuv>`
	`<tuv xml:lang="Lithuanian">`
	`<seg>`Ši Galaktikos zona buvo man nežinoma, nepažinojau nė vieno žvaigždyno. Už siauro langelio mirguliavo žėrinčios dulkės.`</seg>`
	`</tuv>`
	`</tu>`

4.2. Bulgarian-Polish-Lithuanian Comparable Corpus

The practice of collecting texts in different languages on the basis of similarity of type, content or function was common in translation research and training before the word "corpus" came to mean a collection of digital texts.

Ever since 1988 the benefits of using texts from printed public bilingual notices, tourist brochures and instruction manuals for translator training were discussed.

Technological advances have now made it easy to construct great repositories of multilingual texts – comparable corpora. A comparable corpus can be defined as a collection of texts composed independently in the respective languages and put together on the basis of similarity of content, domain and communicative function. Comparable corpora can be created from a variety of sources: collections of texts distributed in electronic format (e.g. newspaper archives on CD-ROM, the Internet, etc.), or even from scanned or typewritten material. Criteria for creating comparable corpora could also be different: in dependence of the homogeneity of texts, or in terms of features such as subject domain, etc. The size of comparable corpora can vary depending on how well they meet these criteria, for example, collections of newspaper articles downloaded from Internet can produce great comparable corpora.

The trilingual Bulgarian-Polish-Lithuanian comparable corpus contains literary works representing mostly modern Bulgarian, Polish, and Lithuanian literature (from the second half of the 20th century), with the text size being comparable in the 3 languages, as well as texts from the electronic media. The latter text type encompasses descriptions in the three languages of the same event. The English text is also included. Such texts are specified as "**parallel descriptions of content**".

5. Applications of Digital Corpora

The aligned parallel corpora are useful for many natural language processing applications: in systems for machine-aided human translation, or for training of software for machine translation. They are prerequisite for contrastive studies or other linguistics research, and can also be used for retrieval of linguistic information, for producing concordances, etc.

Training of translators (humans or programming tools)
The main application area of the corpora is the translation.

The parallel and aligned corpora are successfully used as language materials for the training of translators, humans or programming tools for machine translation, as well as in education – for language learning in schools and universities. Instructors, students and professional translators can use bilingual corpora as a complementary resource in educational process. Bilingual corpora can also be used for training of software packages for automatic disambiguation of morphosyntactic annotation.

The advantage of processing a text corpus is to obtain context specific information about syntactic structures and usage of words in a given language. In the case of parallel corpora, one can obtain context-specific correlations between these languages, which are usually much less ambiguous than general collections. Resulting data from these corpus analysis processes can be used to develop context-specific tools for translation and to standardize the usage of structures and word sets for future multilingual document production. This approach is more correct as there is no "word"-to-"word" comparison, but a comparison of word-forms in a broader context, which allows a better identification of the word's meaning.

Development of multilingual lexical databases and digital dictionaries
The parallel and aligned corpora are the best resource for the development of bi- and multilingual lexical databases and different kinds of digital dictionaries.

Multilingual parallel corpora represent a good base of data for bilingual dictionaries creation. There are many research projects for automatic extraction of bilingual lexical knowledge from parallel corpora in the field of information retrieval from large scale text corpora. Thus parallel corpora are successfully used for automatic lexicon extraction.

There one could find and extract many examples of the usage of the words from the corpus in a wide context.

In contrastive studies

The three languages – Bulgarian, Polish, and Lithuanian - exhibit some specific features, occurring repeatedly in several categories. At first, different orthography traditions – the corpora are dataset of written languages and the orthography forms an inseparable part of language analysis.

A significant feature is the analytic character of Bulgarian. In the process of evolution of Bulgarian from a synthetic, inflectional language, to an analytic, flectional language, case forms were replaced with combinations of different prepositions with a common case form. Bulgarian has lost most of the traditional Slavic case system, exhibits several linguistic innovations in comparison to the other Slavic languages (a rich system of verbal forms, a definite article), and has a grammatical structure closer to English or the Neo-Latin languages than Polish. One of the most important grammatical characteristics of the Bulgarian language which sets it apart from the rest of the Slavic languages is the existence of a definite article. Polish lacks the definiteness attribute altogether. The definite article in Bulgarian is postpositive, whereas in Lithuanian a similar function is served by qualitative adjectives and adjectival participial forms, both with pronominal declension. Bulgarian preserves some vestiges of case forms in the pronoun system. Polish and Lithuanian exhibit all features of synthetic languages (a very rich case paradigm for nouns). Bulgarian and Lithuanian have a high number of verbal forms, but Polish has reduced most of the forms for past tense. Both Polish and Bulgarian have a strongly developed category of verbal aspect. In Lithuanian the verb can have more than one aspect depending on the usage of a base stem for present, past and future tense. Some examples, shown in [5], demonstrate well that the trilingual corpus will be useful to linguists-researchers for research purposes alike, for instance in contrastive studies of the three languages together or in pairs.

Concordances

One of the major developments in linguistic research has come from the possibility of studying vast amounts of text through computer

enhanced tools, particularly through text retrieval and concordancing programs. The basic investigation procedure for querying text corpora consists in producing multiple concordance lines, for a specified string of characters — a word, a lemma or a phrase. The citations thus obtained can be sorted to reveal recurring clusters of words. The analysis of these recurring patterns highlights the behaviour of actual language in context, and complements and sometimes challenges the information provided by standard reference tools such as dictionaries and grammars. This means that the opportunity exists – thanks to the broader context – to study more precisely the meanings of word-forms in each language.

Table 5. Bulgakov's *Master and Margarita*: concordance with the Bulgarian word "литераторите"

| **BG**: Кайсиевият сок вдигна обилна жълта пяна и наоколо замириса на бръснарница. **Литераторите** го изпиха и веднага се разхълцаха, платиха и седнаха на една пейка с лице към езерцето и с гръб към Бронная. | **PL**: Morelowy napój wyprodukował obfitą żółtą pianę i w powietrzu zapachniało wodą fryzjerską. **Literaci** wypili, natychmiast dostali czkawki, zapłacili i zasiedli na ławce zwróceni twarzami do stawu, a plecami do Bronnej. | **LT**: Abrikosų gėrimas suputojo geltona puta, ir oras pakvipo kirpykla. **Literatai** atsigėrę tučtuojau ėmė žagsėti, užsimokėjo ir susėdo ant suolelio veidais į tvenkinį ir nugaromis į Bronaja gatvę. | **RU**: Абрикосовая дала обильную желтую пену, и в воздухе запахло парикмахерской. Напившись, **литераторы** немедленно начали икать, расплатились и уселись на скамейке лицом к пруду и спиной к Бронной.

(*Часть 1, Глава 1 „Никогда не разговаривайте с неизвестными" //Интернет-библиотека Алексея Комарова - http://ilibrary.ru //*) |

6. Conclusion

Parallel corpora are the most effective means for the creation of bi- and multilingual dictionaries and contrastive grammars. One has to remember that parallel corpora comprise direct material for the evaluation of translations and their analysis will bring out the improvement of the quality of both traditional, human translation, and machine translation.

Besides, texts extracted from parallel or aligned corpora prove the necessity of evaluating translations: it is common that in translation words get omitted or word meanings get changed. That is why online free-use parallel texts are a useful education resource.

Bibliography
1. Dimitrova, L., Erjavec, T., Ide, N., Kaalep, H.-J., Petkevic, V., and Tufis, D. Multext-East: Parallel and Comparable Corpora and Lexicons for Six Central and Eastern European Languages. In: Proceedings of COLING-ACL '98. Montréal, Québec, Canada, 1998, 315-319.
2. Dimitrova, L., Koseska, V. Bulgarian-Polish Corpus. J. Cognitive Studies/ Études Cognitives. Vol. 9, SOW, Warsaw, 2009, 133–141.
3. Dimitrova, Ludmila, Koseska, Violetta, Dutsova, Ralitsa, Panova. Rumyana. Bulgarian-Polish online Dictionary – Design and Development. In: Koseska, Dimitrova, Roszko (Editors, 2009). Representing Semantics in Digital Lexicography. Proceedings of the MONDILEX Fourth Open Workshop, 29 June – 1 July 2009, Warsaw. SOW, Warsaw, 2009, 76–88.
4. Dimitrova, L., Koseska, V., Roszko, D., Roszko, R. Bulgarian-Polish-Lithuanian Corpus–Current Development. In: Proceedings of the International Workshop "Multilingual resources, technologies and evaluation for Central and Eastern European languages" in conjunction with International Conference Recent Advance in NPL'2009. Borovec, Bulgaria, 17 September 2009. INCOMA Ltd., Bulgaria, 2009, 1-8.
5. Dimitrova, L., Koseska-Toszewa, V., Roszko, D., Roszko, R. Application of Multilingual Corpus in Contrastive Studies (on the example of the Bulgarian-Polish-Lithuanian Parallel Corpus. J. Cognitive Studies/Études Cognitives. Vol. 10, SOW, Warsaw, 2010, 217-240.
6. Dimitrova, L., Panova, R., Dutsova, R. Lexical Database of the Experimental Bulgarian-Polish online Dictionary. In: Garabík, Radovan (Ed.) Metalanguage and Encoding Scheme Design for Digital Lexicography. Proceedings of the MONDILEX Third Open Workshop, Bratislava, Slovak Republic, 15–16 April 2009. Tribun, Brno, 2009, 36-47.
7. Ide, Nancy. Corpus Encoding Standard: SGML guidelines for encoding linguistic corpora. Proceedings of the First International Conference on Language Resources and Evaluation, LREC'98, Granada, 1998, pages 463-470, ELRA. 1998.
8. Ide, N., Bonhomme, P., and Romary, L. XCES: An XMLbased Encoding Standard for Linguistic Corpora. Proceedings of the Second International Language Resources and Evaluation Conference. Paris: ELRA, 2000, 825-830.
9. Leech, G. (2004). Developing Linguistic Corpora: a Guide to Good Practice Adding Linguistic Annotation. http://ahds.ac.uk/guides/linguistic-corpora/ chapter2. htm

Ludmila Petrova Dimitrova, Prof., PhD
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Tel. Office: 02-979-28-33
Mobile: 0887-103-810

STEREOSCOPIC IMAGES.
METHODS FOR CREATION AND VISUALIZATION

Plamen Geraskov, Nikolay Lyutov

Abstract: Stereoscopic images are more widely spread in the field of scientific knowledge, computer technology, design and lifestyle. The advent of fast computers and software allows for creating an ever closer to the real and stereoscopic 3-D models and conversions. The main methods for creating and visualization, which are realized in practical applications: Method using color filters, polarizing filters method using one or more screened devices interfering method filters, optical switching method, Method using a stereoscopic systems autostereoscopic method Holographic methods, Volumetric 3-D methods.

Keywords: stereographs, 3-D, polarization optics, anaglyph, interference, disparity, binocular vision, holography, stereo pair.

Model the process of perceiving a stereo image

Each site has different signs of depth - overlapping objects on the stage, degree of compaction of vitreous to focus, due to the tension of the eye muscles, triangulation and size of the projected image on the retina.

These conditions are met Stereogram - two images taken by two cameras located at a distance corresponding to the distance between the eyes.

It may be an artificial effect of depth for objects which are set by disparity.

To remove other signs of depth using the effects of visual noise created by stereogram with random points.

Method for creating stereogram with random points
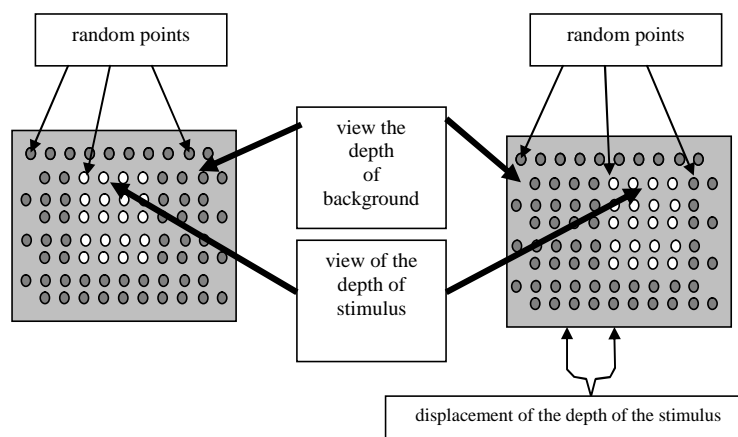
Fig. 1. With random dots stereogram

164

The left image is created by uniformly locating points of a certain color on a background of another color. In this image set view of the depth of background and a separate area - view of the depth of the stimulus.

For the right image is copied to the left image, but the depth of field stimulus moves horizontally at a distance, creating disparity.

Depending on the deviation of this disparity in left or right direction on the stimulus is seen in front or behind the plane of the image.

For each area of Stereogram can be set not only a disparity but disparity group that is perceived by the eye, as ordered translucent surfaces.
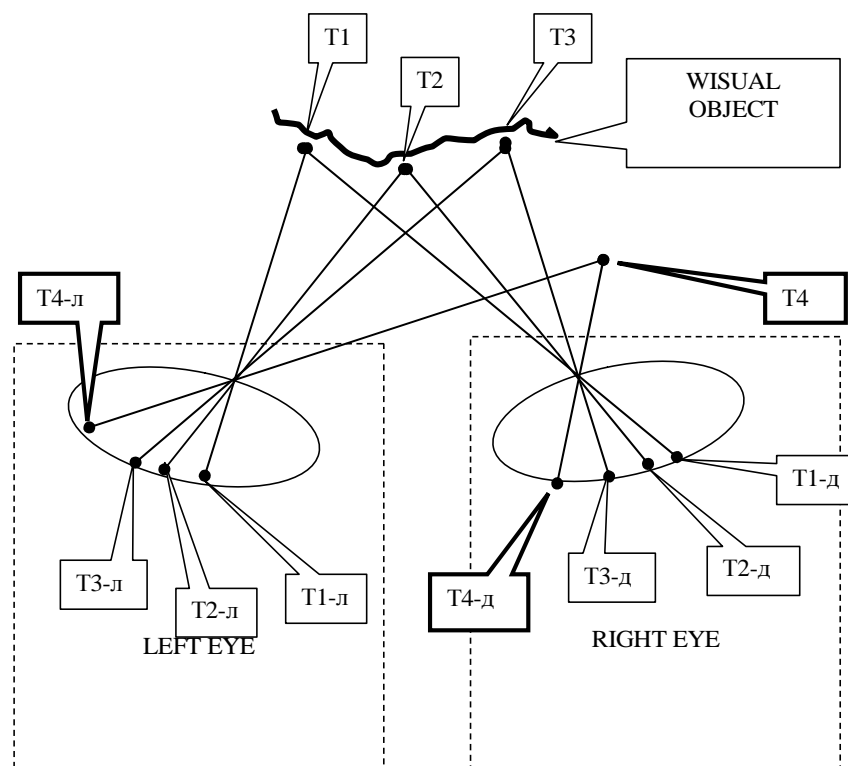
Fig. 2. Disparity

Surface T1 / T2 / T3. The point of fixation - T2 projected onto the center of the "fovea" of the observer.

Projection of the object T1-T2-T3 is designed to (T1-l / l-T2 / T3-l) of the left eye and (T1-l / l-T2 / T3-l) of the right eye.

Point T-4 is the off-site that focuses on the eyes of the beholder. And disparities points projected onto the retina. In the example, T-4 is

closer than the fixation point of T-2 and positioned so that projection (T-4, l) is greater than the projection (T-4-e). The location of these points in space to create an image map (network) disparity.

Binocular neurons have different effects on the intensity of the cells of the retina.

The main identified to date, adjusted for adoption of disparity neurons are:

- Inflammatory - "tuned-excitatory" - neurons respond to a maximum intensity of reactions to stimuli that disparity which are set.

- Restraint - "tuned-inhibitory" - neurons respond to minimum intensity of reactions to stimuli that disparity which are set.

- For distant stimuli - "far" - neurons respond to a maximum intensity of effects of incentives, located farther from the point of fixation - T-2

- For close irritation - "near" - neurons respond to a maximum intensity of effects of incentives, located closer than the point of fixation - T-2

In the human eye center binocular neurons are about 84%, of which about 50% respond with random dots stereogram.

Three-dimensional perception of reality environment is possible as a result of the presence of binocular vision. It is formed by two forward-facing eyes.

Both eyes are located at a distance and thus sent to the brain, two images are shifted relative to each other. This shift is called parallax. It is the reason for the ability to perceive the image in depth.

All methods for creating stereo effects are based on the principle of creating two images to be displayed separately on each eye.

Method with colored filters

It was created by Wilhelm Rollmann, 1853. Common name anaglyph method for forming a visual stereo from simple two-dimensional images through color coding of images, respectively right and left eye of the beholder.

The term "anaglyph" comes from the Greek "anáglyphos" - relief. Monitoring is done through glasses with color filters of opposite colors, mostly blue and red.

Stereoscopic effect is achieved by overlapping the two images respectively having opposite color tones.

Since then, numerous applications were made of anaglyph stereo images, both in static images - photos, topographic maps and in the field of dynamic images, three-D cinema.

The main advantage is the low price of used glasses with color filters, a very rapid adjustment of the eye.

No need for special additional equipment when using a computer screen, no flicker and blink.

The main disadvantage is narrowed colors and discomfort when watching dynamic video images.

Other disadvantages are:

- The need to precisely match the color characteristics of image display and blue - red filters of the glasses. If this condition is not met is obtained contours split;

- Loss of stereo effect on objects with colors close to blue and red;

- Temporary loss of color vision after prolonged viewing of anaglyph images and discomfort.
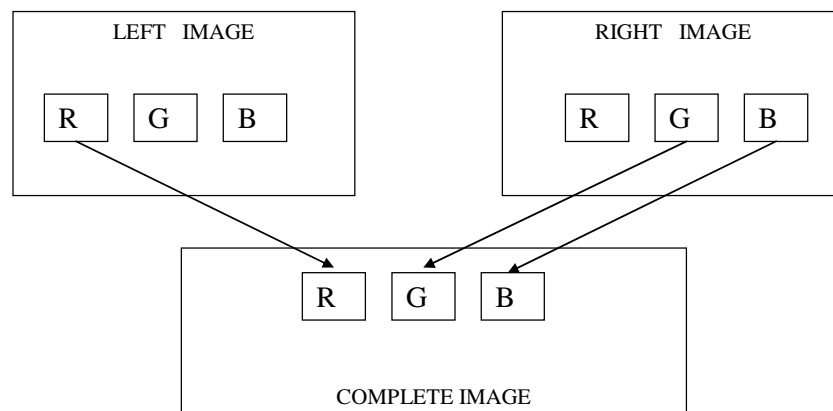
Fig. 3. Anaglyph

Method with polarizing filters

Use the principles of linear and circular polarization.

In the method of linear polarization, two images are superimposed over one another on a single screen.

The two images are projected from two synchronized video projectors in linear polarized filters. Polarization vectors of one filter is rotated 90 degrees to the other filter. And only 0.71 on the strength of light falling on the screen.

The screen must be able to prevent depolarization and loss of brightness. This screen is used with silver or aluminum coating.

Use passive linear polarized glasses. Polarizing filter for the eyes are also rotated relative to one another by 90 degrees.

A disadvantage of the method of separation of images for both video channels with linear polarization is the loss of stereo effect by tilting the head of the viewer. Typical of this principle is IMAX-3D.

In the method of circular polarization, two images are superimposed on one screen, but in this case left and right video channel have opposite directions of rotation.

The illusion of a stereo image is achieved by observing the image through glasses with filters, which are also circular polarization. In this method of dividing the video channel is not required the viewer to keep his head upright.

Typical of this principle are MasterImage and RealD Cinema.
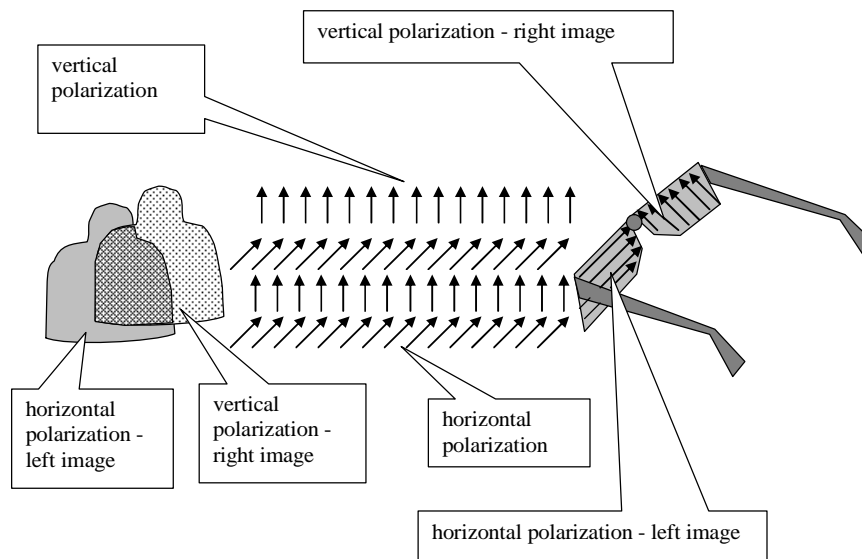
Fig. 4. division of video channels by linear polarization

This method ensures accurate display of colors and good detail. Used glasses are passive, light and cheap. Display working in dual mono and stereo mode resolution, dual lower resolution, which is easily switched.

Option splitting of the optical channels using different polarization displays are based on method "interlace". These displays simultaneously output two images, one is formed by the odd lines and one of the even lines.

168

The polarization of the odd lines is opposite to the polarization of odd. In observation of such images in polarized glasses in the human brain forms a complex stereo image.

This principle of locating the two interlaced images is reduced twice the vertical resolution. Particularly troublesome effect of the vertical tear occurs in text mode.

Method for establishing a polarized pair of images with two screens and translucent mirror

Both displays are angled one above the other. Each shows an image that is polarized opposite to the other. Half mirror is located on the bisector.

It made the reconciliation of the two images - lost from the lower display and reflected from the upper display.

With appropriate adjustment of the brightness of both displays two images are aligned in brightness and create a complete stereo pair without loss of resolution in vertical and horizontal, and without blinking.

Disadvantages of this method are:

- Limited by the mirror field, making it applicable only to an observer;

- Large dimensions;

- Easy loss to reconcile the image with minimal deviations in the position of components and separation of the image.
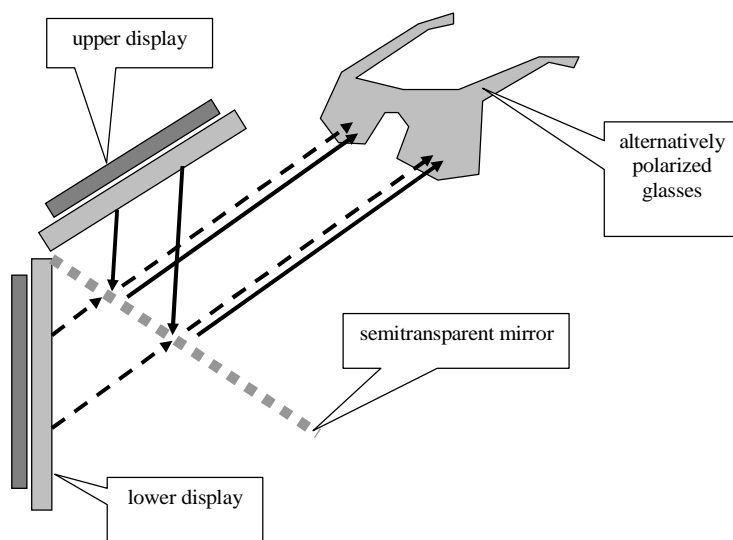
Fig. 5. Creating a stereo image in a mirror and two translucent display

There are variations of the scheme in which the lower display is horizontal and the top is vertical.

Method of interfering filters

In this technology each eye forms an image with different wavelengths. Accordingly, for red, green and blue. Glasses with this method filter corresponding wavelengths and thus two separate video channels for each eye.

The method requires very precise balancing of the display colors to comply with the model color filter glasses.

The disadvantage of this method is the partial loss of the false colors and displaying certain colors.

Method using an optical switch

The idea was proposed by Almeida in 1858. Is implemented by banks in 1936.

In this method the left and right image is formed one after another sequentially across the screen.

Monitoring is done by eye with fast aperture, which covers the eyes consistently turn in one and the other eye - "eclipsing" glasses. Were originally used glasses with mechanical switching of opaque blinds.

Nowadays, using optical plates with fast liquid crystals. This switching of the light beam between one and the other eye is synchronized with shifting images. As a result, each eye sees only intended to give the image.

Thus, on account of the effect of inertia in vision creates the illusion of continuous two different video streams. Of these, three-dimensional image is formed.

To ensure flicker-free image should be used with dual displays under high frame rates - at least 120 Hz. Each eye sees its image with a frequency of 60 Hz.

With the same frequency shift serial and aperture of the glasses. Eclipse method of creating stereo images is suitable for use with LCD - displays where there is no winking and giving a more stable image than CRT - displays.

On the other hand the requirements for LCD - displays are high - frequency is needed to update the frame at least 100 Hz.

Synchronisation between the display and glasses is remote, via infrared or Bluetooth technology. Despite relatively high frame rate, the effect of flashing is available.

This disadvantage can be overcome partially by increasing the refresh rate of frame - 240 Hz.

Disadvantages of this method are:
- Lower brightness compared to mono displays;
- Heavier glasses containing electronic components and power;
- The higher price of the glasses;
- Occurrence of headaches and eye fatigue;
- Discomfort when using spectacles.

Not overcome the disadvantage - loss of stereo-speed action shots, this manifests itself in a split image.

The main disadvantage of method using optical switching Flicker is due to the low performance of the existing liquid crystal aperture of the glasses.

The aperture opens with considerable delay to the new frame. At a frequency of frame rates - 120 Hz, the active time frame for each channel - 16 microseconds are displayed as a visible representation only 2 microseconds.

During the rest of the aperture is closed and the corresponding glasses is opaque, causing noticeable blip. Moreover, too little time in which the aperture of the glasses are opened sharply lower the brightness of the image that extra strain on the eyes.

Method using a stereoscopic systems

Method using a stereoscopic systems.

The first patent for such a system is brought by William Grin 1890.

The method is a projection of two separate images on two horizontally arranged side by side panels.

Surveillance by stereoscope, allowing each eye to only see one image. In application of the method over the years were used as two synchronized photo strips and common band with two images.

The system Space-Vision 3D stereoscopic pair of images are located along the bar above each other. Separation towards both screens is performed by mechanical optical system.

The system "Stereovision", stereoscopic pair of images are arranged as parallel imaging, bent horizontally. With anamorphic projection lenses assure the extend them to restore normal size.

Both systems require the use of special photographic and projected optical-mechanical system. This predetermines its relatively low prevalence.

Stereoscopic systems experiencing a renaissance with the advent of digital stereo.

Stereoscopic systems are built on binocular vision and can be classified as a full version of the demonstration of volumetric images.

But when changing the angle and motion in space is available loss of stereo image of reality, occurred because the difference in the views of the shooting stereo camera and viewer.

Stereoscopic displays:

Stereo images are separated by a plane perpendicular to the screen allowing the appropriate eye to see only intended to give the image.

RIGHT IMAGE

LEFT IMAGE

SEPARATING PLANE
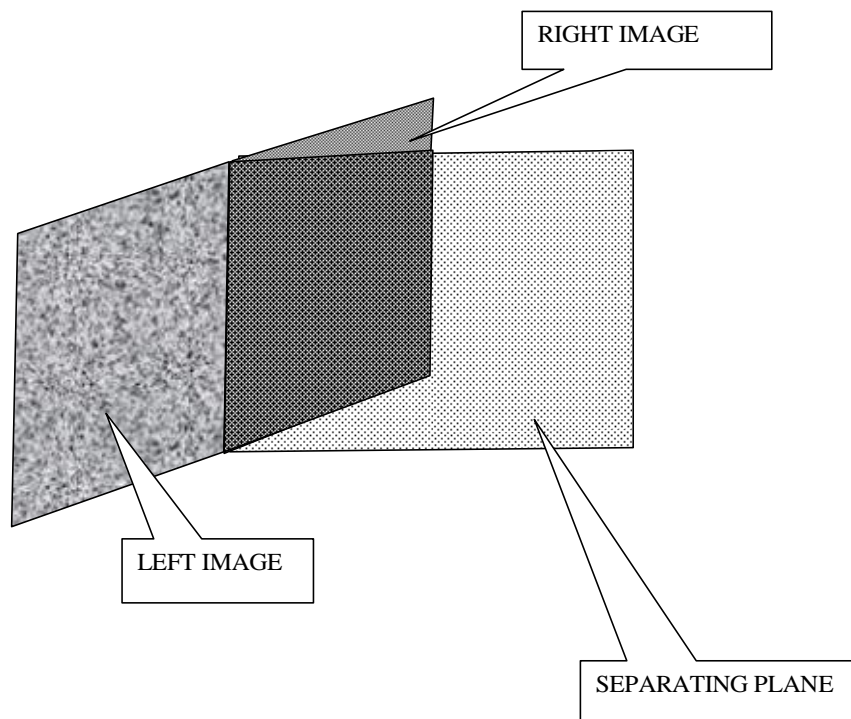
Fig. 6. Separation of two video channels

Stereo image by parallax barrier screen grid

The separation of stereo views shall be carried out through vertical parallax grid. In a typical variant to LCD-monitor is placed a grid of opaque bands.

Each strip has a width of two pixels. The distance between the opaque film is one pixel. In the right angle and distance from screen to the left eye sees only even pixels and odd right only.

In modern versions vertical parallax lattice is created by an additional panel on the basis of liquid crystal, which creates an additional opportunity to exclude the stereo mode, and monitoring of mono images.
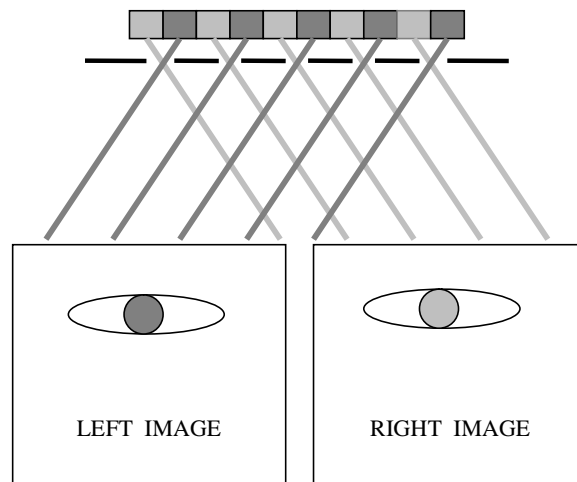
Fig. 7. parallax grid

Stereo image by parallax surface of vertical cylindrical lenses - optical Fresnel filter. On the basis of internal refraction of light rays from each pixel, this principle deviates video information from the even pixels in the zone of the left eye and odd to the right.
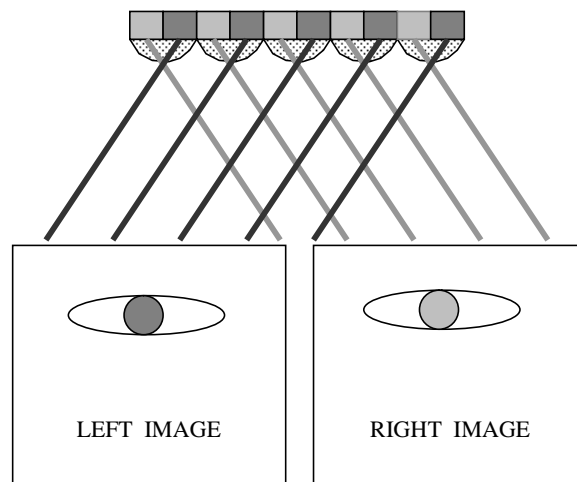
Fig. 8. Cylindrical Lenses

Disadvantages of both methods are:

- Creating more than one definition for the stereo plane "left - right" instead of giving rise to a series of panels, with a number equal to half the number of pixels horizontally;

- When moving the viewer from the central plane parasitic stereo arise where part of the video image intended for one eye falls on the other;

- Halving the horizontal resolution;

- Inability to sighting of the object and dynamic parallax;

- Very limited area for a real stereo, which corresponds to the capture of stereo camera.

Auto stereoscopic method

It contains of auto stereoscopic 3D-displays that appear three-dimensional image without the use of additional accessories for the eyes.

Use spatial separation of the stereo pair through the parallax barrier, which is a second liquid crystal plate. This additional component appears in optical system included a network of narrow vertical opaque strips.

The main screen provides a sequence of images for left and right eye. This is done by fogged lamplight consistently for one and the other vertical columns of pixels.

In front is a raster image with the same step across, which at sufficient distance viewer provides two separate images for each eye of the beholder.

The main advantage of this method is the elimination of the need for glasses for the observer.

The method has development systems, tracking the position of the viewer, through infrared and ultrasonic sensors.

But these systems to identify the location of the viewer only work where the viewer is one.

Some of the systems implemented this principle instead of finishing arrays using vertical cylindrical Fresnel lenses.

This avoids the total reduction in image brightness at the expense of residual illumination of the Fresnel lens.

Disadvantages of Auto stereoscopic method for displaying stereo images are:

- Double the lower horizontal resolution over the loss of 3D effect, if the observer is outside the central angle;

- Parasitic optical effects, such as "arc", "muar", skip of views;

- A general decline in brightness of the image;

174

The disadvantage of the method using columns of cylindrical Fresnel lenses, compared with the method using a parallax barrier is the inability to switch to mono mode.

light bulb

linear vertical raster for left eye

parallax barrier

aperture grid

linear vertical raster for right eye
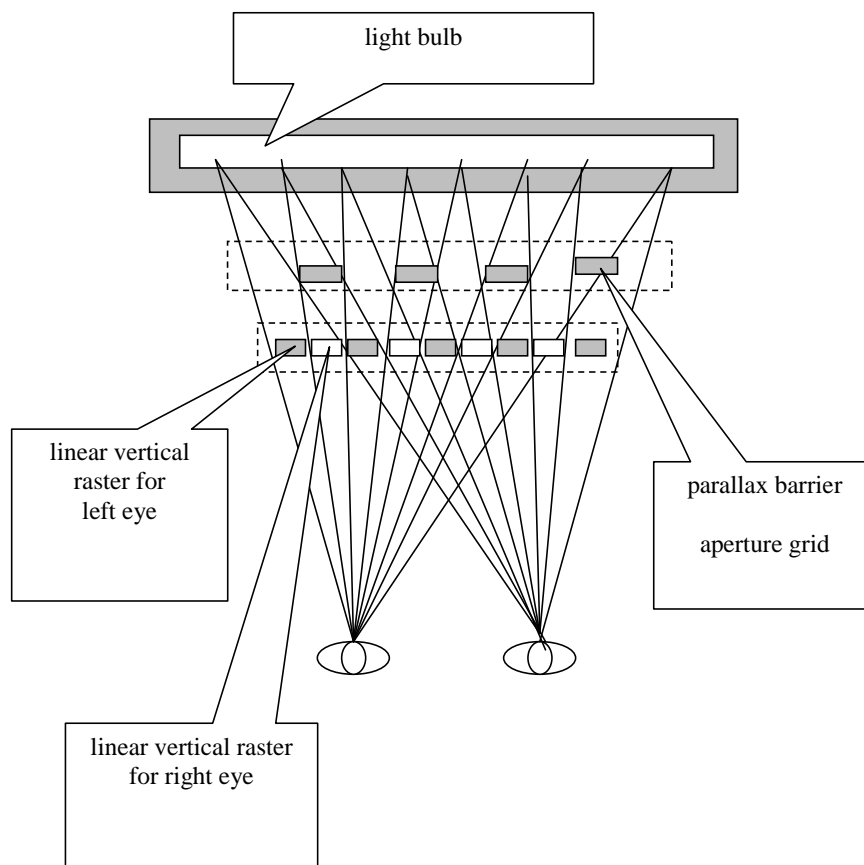
Fig. 9. Auto stereoscopic method using LCD - Screen

Holographic methods

Using 3D-holographic displays. Dimensional image formed by simulated spatial distribution of light waves as they are located in the reflection of the real object. This is achieved through the visualization of spatial pixels - "Voxels" in three-dimensional matrix is displayed when a static hologram and "Doxsels" - varying in time "Voxels" in dynamic holograms. Holographic methods for creating dynamic images are on stage - laboratory tests.

Volumetric 3-D methods

Volume using 3D-displays. Dimensional image formed by the placement of pixels within a given volume carried by mechanically synchronized three-dimensional rasterization, which is accomplished by vibrating or rotating reflective elements located throughout the volume which has three-dimensional image. The rate of coverage of each element of the surface is high enough to use the visual effect of inertia of the observer and also to form an overall perception of the surface. At present only laboratories have developed versions of these methods.

References
1. A. Jean Etienne Gaudreau, Mark Bechamp Vince Power "Innovative Stereoscopic Display using variable polarized angle"
2. O. Tishutin, T. Striegler "Stereoscopic 3D and iZ3D Perception"
3. Sergei Knigin "3D display" http://www.3dnews.ru/display/3dd/
4. www.neurokoptics.com
5. www.stereo-pixel.ru/lcreflex.htm
6. www.zalman.co.kr/ENG/

Authors:
Assistant, Eng. **Plamen Geraskov Ivanov**, Chernorizets Hrabar Free University of Varna, Faculty - International Economics and Administration, Department of Computer Studies, email: plamengeraskov@abv.bg, Tel.: +359-31-21-82
Professor **Nikolay Rashkov Lyutov**, PhD, Chernorizets Hrabar Free University of Varna, Faculty - International Economics and Administration, Department of Computer Studies, email: lyutov@vfu.bg, Tel.: +359-52-359-571

DEVELOPING MOBILE APPLICATIONS WITH PHONEGAP

Krasimir Ivanov

Abstract: *PhoneGap is an open source framework for developing mobile applications. It allows the use of HTML5, JavaScript and CSS3 for developing. It supports major mobile operating systems and allows web developers to start creating mobile web applications. It can be used to develop client applications, to access the database driven sites or services. Modern JavaScript libraries can be used to facilitate the work of the programmer.*
Key words: mobile applications, Phonegap, HTML 5, JavaScript, CSS3, Android, iOS, Windows Mobile, Blackberry, Ajax, REST, JSON

Goal: Introduction to the new technology for developing mobile applications based on Web standards.

The main problems faced by developers of applications for mobile devices are the huge differences between the devices. Firstly, big amount of different operating systems, followed by differences in hardware, display size, programming languages for the different operating systems.

The languages used to create applications for different operating systems are:
- Windows Phone 7,5: C # and HTML JavaScript rather
- Android: Java
- iOS: Objective-C
- Bada: C++
- Symbian: C++ with QT
- WebOS: C/C ++ or HTML5, JavaScript, CSS with Mojo or Enyo
- Blackberry: Java

In addition, each manufacturer of mobile devices built-in them different sets of additional components such as: GPS, compass, gyroscope, camera, touch screens and more and has its own methods for accessing and managing these additional components. This makes it very difficult to develop mobile applications for different operating systems and devices. Difficulties for developers in some cases come to the point that in order to develop application for a device requires a specific operating system on the workplace of the programmer and a physical presence of a mobile device. In most cases this can be avoided with the use of simulators for different operating systems. These difficulties are partly avoided by using technologies like PhoneGap.[R5]

PhoneGap is an open source framework[1] created by Nitobi Software and purchased by Adobe on 04/10/2011. PhoneGap provides

---

1   A set of functions that allow access to device hardware.

unified access to different API for different devices allowing to create the logic of an applications using JavaScript, HTML 5 and CSS. To facilitate writing JavaScript it can be used various libraries such as jQuery, jQuery Mobile, Sencha Touch and others. One advantage of PhoneGap is that it uses standardized technologies that people already know. The only difficulty may be somewhat JavaScript, but it can be learned more easily through the various libraries and frameworks. Another important advantage is that when an application is written once it can be packaged and work for various operating systems.

Applications created with PhoneGap, are hybrids. They are neither own nor purely web based. All of the visualizations are made by "web view" instead of Objective-C in Android. Many of the features of HTML 5 are supported by the browsers in mobile devices and this facilitates the development of the applications. Access to the hardware components of the device via the API is provided by the PhoneGap. The disadvantages are that the hybrid applications don't have absolutely full access to all hardware components.

PhoneGap run an application written in HTML 5, JavaScript and CSS in the browser's window of the device. By default, the code written in JavaScript can not access the hardware components of the device for security reasons, but PhoneGap provides the ability to access these hardware and software components (list of contacts, pictures, files, sms, etc..) in a uniform manner by objects and methods in JavaScript. From the fact that the functions of the operating system and hardware are not called directly, but through PhoneGap, which converts queries sent via JavaScript in requests to specific hardware components occurs slight delay. Information about the performance you can see it bellow in [Test 1]. The average delay is about 3 times, but this will improve with the release of each new version of the product [Test 1].

Development tools:

There are two options for application development. The first one is using IDE according to the programmer choose and the other one is DreamWeaver CS 6 or newer.

For the first option is needed:

• Installing a programming environment such as: Eclipse, Xcode, Visual Studio

• Installing of a SDK for the OS.

• Installing PhoneGap

If using DreamWeaver:

• Creating a site

178

- Installing a SDK for the specific OS.
- Choosing an option - mobile application
- Assembling and installing the application on the device.

In all cases it is necessary the availability of the appropriate SDK for the various operating systems.

How to develop apps with PhoneGap:

PhoneGap is only a library that is added to the application along with some JavaScript and xml files. This allows communication to objects and methods which are defined in it to give access to the hardware of the device. All of the program application logic is written using object-oriented programming on JavaScript. The writing process has to be done by the standards for HTML 5 which gives the programmer the opportunity to use its innovations such as structural tags for page content, local storage, animate objects and effects to move from page to page, audio, video tags and more. Thanks to JavaScript libraries it works easily with AJAX regardless of a browser of the device. This allows the application to take advantage of a server logic to carry out part of the work of the application request and returns the result in the form of JSON. Thus you can easily work with different web services, which provide different services. Thanks to CSS 3 you can work with animated effects and transformation of images.

Essentially, to start developing a mobile application with PhoneGap are needed the following things: install an appropriate SDK for the operating system, IDE for the OS, install PhoneGap. During installation of these things they are integrated and development can begin. In essence, the development of application is very similar to developing a web site. The difference is that everything must be consistent with the design and the idea for working for mobile devices with many different screen resolutions and with limited memory. Therefore it's good all of the resources for the application such as images, css, javascript to be as small as possible. It should also be provided a use of local copies of external libraries instead of loading them from the Internet because the application can work when the device is not connected to the Internet. Also when writing a JavaScript functions the programmer should be careful and release resources that are no longer used.

What is doing PhoneGap:

It creates a window in your browser and starts the application in it. Thanks to some xml and jar/dll files an access to the device's own API is provided.

179

Workflow of PhoneGap mobile application development

Createon of an application based on web standards.

• Writing according to the HTML 5 standard code.

• Packeding with PhoneGap.

• Usage of open source framework or "PhoneGap build"

• Usage of built-in API to access the hardware components of the device.

Installing on multiple platforms

PhoneGap build is a cloud service provided by Adobe to enable developers to package their applications for all supported platforms at once. The service is free if the applications are open source.

Advantages and disadvantages

Advantages: The programmer write the code in HTML 5, CSS and JavaScript, which are well known. Same code can be runed on all platforms. It is possible to be used the systems for distribution and payment such as Android Market or AppStore.

Disadvantages: Performance is lower if you make heavy mathematical calculations for games or other applications. It is possible to write some parts of the application on the language of the platform to improve it's performance by caching some of the information and using the hardware acceleration for the graphic objects. There are no many prefabricated elements for user interface "widgets", transition effects and standard controls. The work may take a little longer.

Application and use

PhoneGap is suitable for creating a small applications when there is no time and enough human resources to create an application that works on many different operating systems.

Undoubtedly PhoneGap is not a universal solution to all of the problems in developing mobile applications. It has its disadvantages. It is slower than applications written in their own programming languages of the respective operating systems, development of applications may take a little longer time, not all operating systems has full support for all components (Table 1).

Phonegap is suitable to create applications that use server-located database, applications which can access to online stores and other sites providing services such as ticketing, weather forecasts, stock information and more.

Writing applications on the programming language of the respective operating system is always the best solution.[R7]

180

Application compatibility information at the time of writing [Table1]

	iOS iPhone / iPhone 3G	iOS iPhone 3GS and newer	Android	OS 5.x	OS 6.0+	WebOS	WP7	Symbian	Bada
ACCELEROMETER	✓	✓	✓	✓	✓	✓	✓	✓	✓
CAMERA	✓	✓	✓	✓	✓	✓	✓	✓	✓
COMPASS	✕	✓	✓	✕	✕	✕	✓	✕	✓
CONTACTS	✓	✓	✓	✓	✓	✕	✓	✓	✓
FILE	✓	✓	✓	✓	✓	✕	✓	✕	✕
GEOLOCATION	✓	✓	✓	✓	✓	✓	✓	✓	✓
MEDIA	✓	✓	✓	✕	✕	✕	✓	✕	✕
NETWORK	✓	✓	✓	✓	✓	✓	✓	✓	✓
NOTIFICATION (ALERT)	✓	✓	✓	✓	✓	✓	✓	✓	✓
NOTIFICATION (SOUND)	✓	✓	✓	✓	✓	✓	✓	✓	✓
NOTIFICATION (VIBRATION)	✓	✓	✓	✓	✓	✓	✓	✓	✓
STORAGE	✓	✓	✓	✓	✓	✓	✓	✓	✕

Table 1. Supported Operating Systems with version 1.7.0

Phonegap is not the only option for developing mobile applications based on Web standards. Another examples are Appcelerator Titanium, M Project, NimbleKit, Sencha Touch, Construct 2 and many others.[R7]

The main disadvantages of other products are their high prices. Some are specialized in developing only games like Construct 2. Others require the usage of their own programming languages. Appcelerator Titanium use JavaScript, but it does absolutely everything like styling elements of the interface, manipulating them and everything is done by calling out functions. This cause difficulties to start developing an applications for people who until then were developing websites. Phonegap can be used completely free and it's not restricted to a specialized activity. Realized Phonegap functionality is the least common denominator of functionality on all supported platforms. This functionality is relatively small, and it's easy to add support for new operating systems. The possibilities of Phonegap can be extended by modules written in programming language for the specific operating system, which improve

the application performance, but requires development of such module for each operating system for which the application should work. The reason that Phonegap is managed by the Apache Foundation and sponsored by Adobe ensures that the product will evolve and improve with each, a new version. At the time of writing it is version 1.7.0, but it is soon to be released 1.8.0 r1.

After analysis with [R1] Test 1: Results of the tests with Phonegap-1.3.0 are the following:

Three different loop times were measured in javascript and native android java. Example Code in Android was:

```
package com.marguspala.perftest;
import android.app.Activity;
import android.os.Bundle;
import android.widget.TextView;

public class PerfTestAndroidActivity extends Activity {
 /** Called when the activity is first created. */
 @Override
 public void onCreate(Bundle savedInstanceState) {
   super.onCreate(savedInstanceState);
   setContentView(R.layout.main);
   String text="1234567890";
   long start = System.currentTimeMillis();
   for(int i = 1; i<1000000;i++){
      text.indexOf("0");
   }
   long end= System.currentTimeMillis();
   TextView tv = (TextView)findViewById(R.id.textView1);
   tv.setText("It took: "+(end-start)+"ms.");
 }
}
```

Phonegap and javascript was:
```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
<title>Insert title here</title>
<script type="text/javascript" charset="utf-8" src="phonegap-1.3.0.js">
```

```
</script>
<script type="text/javascript">
 function calculate() {
  var text = "1234567890";
  var start = new Date().getTime();
  for ( var i = 0; i < 1000000; i++) {
    text.indexOf("0");
  }
  var end = new Date().getTime();
  var div = document.getElementById("text");
  div.innerHTML = "It took: " + (end - start) + "ms";
 }
</script>
</head>
<body>
 <div id="text">Hello phonegap</div>
 <a href="#" onclick="calculate()">calculate</a>
</body>
</html>
```

It turns out that javascript is roughly 3 times slower than android 2.3.5

100 000 cycles:

Android: 0,1s

PhoneGap: 0,3s

1 000 000 cycles:

Android: 0,8s

PhoneGap: 3s

10000000 cycles:

Android: 7,5s

PhoneGap: 26s

References

1. https://github.com/GameMaker/PhoneGap-Performance-Test--PGPT-
2. http://marguspala.com/phonegap-performance-measurement-results/
3. http://en.wikipedia.org/wiki/Multiple_phone_web_based_application_framework
4. http://www.scirra.com/blog/85/the-great-html5-mobile-gaming-performance-comparison
5. John M. Wargo, PhoneGap Essentials: Building Cross-Platform Mobile Apps, Adison Wesley Professional, 2012.
6. Matthew David, HTML5 Mobile Websites: Turbocharging HTML5 with jQuery Mobile, Sencha Touch, and Other Frameworks, Focal Press, 2011.
7. Rohit Ghatol, Beginning PhoneGap: Mobile Web Framework for JavaScript and HTML5, Apress, 2012.

Author:

Assistant engineer **Krasimir Todorov Ivanov**, Varna Free University, 0887380054, krasimir19@yahoo.com

METHOD FOR MEASURING THE MUTUAL MOVEMENT OF ELEMENTS OF THE DEVICE FOR SEA WAVES ENERGY EXTRACTING

Vladimir Ivanov

Abstract *This paper presents the method for measuring the mutual relative motion of the vertical surfaces of the device intended to convert the horizontal movements of surface layer of sea waves energy into mechanical one.*
Keywords: *renewable energy, sea wave, Matlab, image processing*

One of unconventional renewable energy sources of the next generation becomes the marine and ocean waves. With their inherent about 800 times denser than air, even in low excitement they can produce significant amounts of energy. To extract this energy it can be used large pontoons, providing vertical movement or turbines installed underwater.

Currently, a project called INWECO processing the capabilities for extracting energy from sea waves becomes a collaborators from the Institute of Oceanology, the Institute of Hydro and Aerodynamics, Central benchtop lab of Mechatronics and Instrumentation of Science and the Technical University of Varna.

Under the restrictions of this project an experimental model of a device for converting energy of horizontal motions in the surface layer of sea water into mechanical energy was developed. The structural scheme of the device is shown in figure 1.
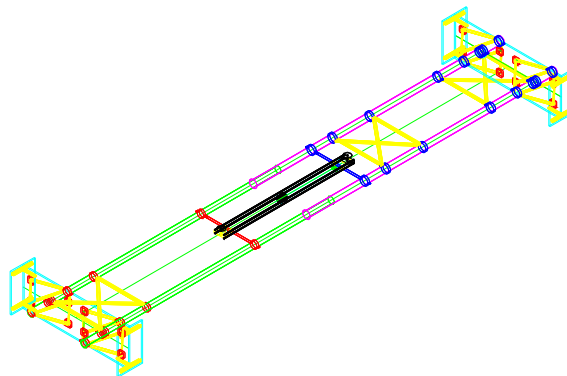
Fig.1. Diagram of experimental device for sea wave energy converting

Essential to the effectiveness of this device are the laws of motion of the two vertical plates located at the ends and in particular the magnitude of the relative speed between them and the individual parameters of wave influence. This problem is dealt with in this article.

To measure the speed of movement of the two vertical plates a two 3W LED were placed of each of them, which in the rest position forms a right quadrilateral. In the process of the device work the distance between two pairs of LEDs and the size of their images are changed. With information derived from these changes can judge the relative size of the speed of the two vertical plates of the device and wave parameters.

These data are obtained from a camera attached rigidly over the device, which transmit images over a time interval. The resulting images are processed in the Matlab enviroument.

According to the theory, developed within the project INWECO for the magnitude of the relative speed of two vertical plates of the device can be judged by the change in the size of the right quadrilateral diagonal formed by two pairs of LEDs.

In order to extract the information about the parameters of the sea wave the fact that the monitoring camera is fixed over the device is using. In this situation due to the impact of sea waves the size of the image of each LED in the number of consecutive frames will be different. This change is used to retrieve information about the amplitude and period of waves. Block scheme of the algorithm is shown in Figure 2.
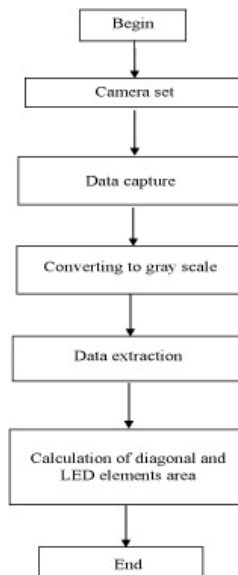
Fig 2. Block diagram of the algorithm

As a set for camera setting the image shown in Figure 3 representing the device at rest is used.The selected camera mode allows easy removal of interest to us objects. In practice, the shape of each object is generally oval, allowing it to be approximated by a circle. This approximation allows to calculate the areas of the images of individual LEDs and thus facilitates the calculation of the statistical characteristics of the object in which the coordinates of their centers of gravity coincide with the geometric center of the circle approximating them.

Thus the calculated coordinates becomes attached to the lower left corner of the image where by default Matlab launches the origin of its coordinate system.[1] Based on the so calculated coordinates of the centers of gravity of individual objects, the length of the diagonal of rectangle is calculated. Under the data sequence contained the areas of individual objects, the period and amplitude of the wave. Can be obtained. The results of all calculations are saved as a vector file to disk.[2,3]
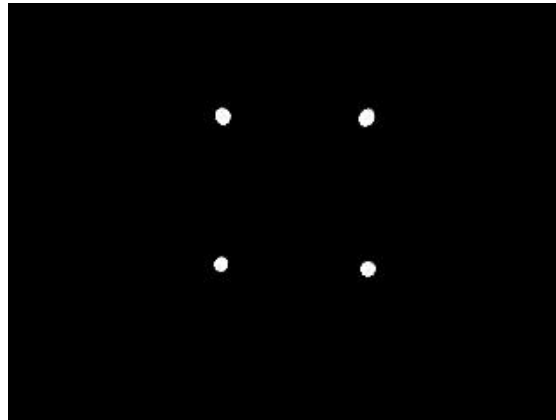
Fig. 3. LED picture used to initial set the camera parameters

A typical image of the spatial position of the LED modeling the impact of waves on the device for converting energy of horizontal motions in the surface water layer into mechanical energy is shown in Figures 4, 5 and 6.

186

Fig. 4

Fig. 5

Fig. 6

The accuracy in calculation of results is within one pixel. Distance behind this pixel in the actual object is determined by the ratio between the length of the field seen by the camera to the number of pixels on the photosensitive matrix of the camera. In this case, the width of the field seen by the camera of 600 mm and 560 pixels across, the real accuracy in this direction is 1.07 mm. Similarly for the vertical accuracy is obtained 1.4 mm.

A certain idea of the amount of diagonal length change derived from the movement of the two vertical plates at the ends of the developed device for converting energy of horizontal motions in the surface water layer into mechanical energy can be derived from the processing of successive frames from a video camera. Model change the size of the diagonal obtained after processing 500 of simulated images is shown in Figure 7.
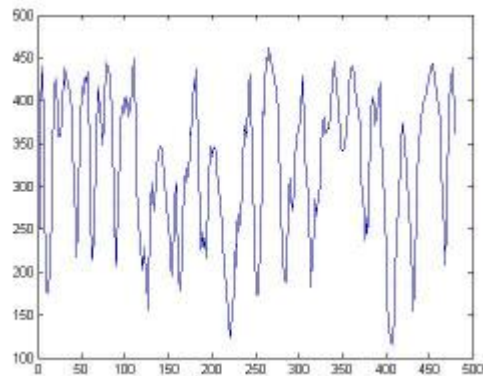
Fig. 7

An idea of the amplitude and period of the wave, resulting from the treatment of 500 simulated images is shown in Figure 8.
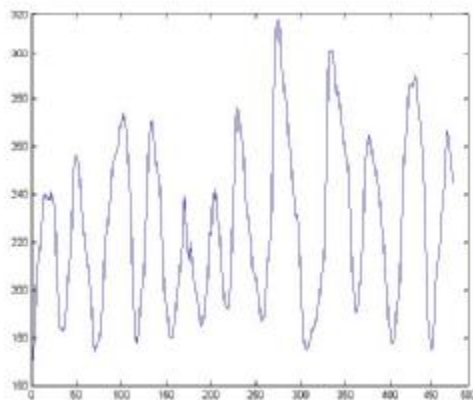
Fig. 8

References
1. Tonchev J. Matlab 7 т.I, "Technika ", Sofia 2007.
2. Tonchev J Matlab 7 т.III, "Technika ", Sofia 2009.
3. www.mathworks.com/nn_image

Author:
PhD-student **Vladimir N. Ivanov**
IICT BAS
Acad. Georgi Bonchev Street bl 2
Sofia 1113
e-mail: ivanov.vladi@gmail.com

188

DESIGN FROM DATA: HOW TO USE REQUIREMENTS FOR BETTER INFORMATION SYSTEM ANALYSIS AND DESIGN

Kalinka Kaloyanova

Abstract: *One of the biggest challenges facing modern complex information systems is providing appropriate methods for their analysis and design, based on real customer needs. Although various methods, approaches and technologies has been presented for analysing and designing IS, no general method exists to explain clearly the design issues and their connection with customer requirements for the software system. In this paper we propose an integration of several methods into a specific design medium, emphasizing on iterative application of different elements of the methods that help creating better system design based on customer data. Some results that support the main concepts implementation are presented, too.*

Key words: *Information System Analysis and Design, Contextual Design (CD), CD Work models, Unified process, Use case Modelling, Domain model*

Research motivation

All approaches for Information Systems (IS) analysis and design should take into account the issues that concern dynamics of modern socio-technical aspects of the systems and to focus on the user role in them. It is obvious that the software should satisfy customer needs but the software industry is full of examples of failed projects that did not deliver what people really needed. On the other hand, the best software product designs happen when the product's designers are involved into the process of collecting and interpreting customer data and appreciate the real need of users [7].

Despite of diversity of analysis and design methods that use modern visualization and communication techniques [9], the role of data, obtained from customers is not well addressed and the mechanism of their transformation into design decisions is not clarified.

In this paper we concentrate on a combination of several methods that focus on better understanding software functionality based on customer view of future system and propose a set of models and techniques that help both analysts and designers to present their ideas and to communicate them with the customers.

Requirement elicitation – methods and techniques

The software requirements are described as "a desired features, properties, or behaviour of a system" by Object Management Group (OMG) [4r]. A lot of methods and techniques for requirement gathering, analysis and management are known. Many definitions and classifi-

cations are summarized in [10]. But in the most cases the diversity of the methodologies and the ambiguity of their description don't help the design team to choose the right method.

On the other hand requirements analysis produces large amount of customer data. This data have to be shared among analytics, designers and developers. Traditional methods like presentations, reports, and other text descriptions don't provide full understanding of all concepts. A good way for cross-functional teams to understand customers need is to apply the methodology of Contextual design [2].

Contextual design

Contextual design (CD) presents a good explanation of the customer's role in software system design. The method provides a set of techniques, models and procedures for step-by-step process that assists the software design in a very systematic way. As it focuses on how the designers can understand the customer needs it is very convenient for IS analysis and design.

Contextual design also presents five models – Flow model, Sequence model, Artifact model, Cultural model, and Physical model. These work models show different perspectives of the system and describe this information on several types of diagrams [3]:

- The Flow model shows the communication and coordination between users;
- The Artifact model captures the physical things created to support the work;
- The Sequence model presents detailed work steps to achieve user intents;
- The Culture model notes the constraints of policy, culture or values;
- The Physical model describes physical structure of the work environment.

According principles of CD the work starts with *Contextual inquiry.* This is the first step of understanding the real customers and their needs. The design team conducts a series of interviews with customers at their workplaces [6].

Usually a lot of different groups work on interviews to understand the functionality of the system. It is important to ensure that all stakeholders get a common description of the system activities. So the next step is *team interpretation session* where every team member use

190

his individual notes from the interviews. As a result a *"whole story"* is unambiguously described in detail using different work models [3].

After all interviews during a *Consolidation session* the data from individual users are consolidated to present a complete picture of the work. The notes from different users are brought together into an affinity diagram, a hierarchical representation of the issues labeled to reflect user needs. Then the team reviews the models and presents the vision – a big picture of what the system could do. In this way all different point of views are consolidated into a coherent system.

CD continues with storyboarding process where work tasks are sketched in as pictures and text. They may include not only system behaviour, but also manual tasks, initial user interface ideas, some business rules. From these storyboards designers build the next element: *User Environment Design (UED)* - a first representation of system functions and how they could be organized in systematic way that corresponds to user goals. The last part of CD is prototyping as a form of discussing user interface with the clients of the future system.

Use case modelling

The use case model is now common form for specifying system functionality, but the model is defined as a part of Requirements discipline of the Unified process (UP).

The *use case modeling* considers a structured way of capturing the behavioral requirements of a software system. Every *use case* defines one specific way of using the system that has being designed. All behaviour of the system is captured by the set of use cases. The users – people and other systems interacting with the system also are part of the model as *actors*. Together the use cases and the actors form the *use-case model* of the system [5].

As about other types of requirements, the UP categorize them into several categories. The acronym FURPS is used to present functional and non-functional requirements (quality requirements): usability, reliability, performance, and supportability of the system [7].

Starting at Inception phase, use cases capture systems system's functionality in step-by step process that is understandable for users but also could be mapped later onto the classes in an object-oriented model of the future system in the next phases of UP – Elaboration and Construction.

The use case model describes not only system behaviour, it defines the boundary of the system and how this system communicates with actors and other systems.

The actors also are important element in use case modelling. An actor may be a kind of user, or role that users can play, or other system. Cockburn [5] distinguishes two type of actors - primary and secondary. The *primary* actor has goals that are fulfilled by system services. The *secondary* one provides services to the system. Larman [7] adds *off stage actors* in order to have all interests identified on the model.

Use cases are easily understood by non-technical people – they are described in text. They are discussed by clients, developers and their managers.

In this way they connect the two sides of the system – the user-oriented, external view of the system and the technical, internal view of the system functionality.

Integration of known models and techniques into a new approach

As use case modeling is now one of the establish methods for requirements specification it could be used during the analysis phase of a project to identify the information system functionality. And the iterative approach of UP could be used, too. Several questions arise here when we try to implement the method in practice.

According to UP principles the Inception phase could be very short [7]. In many cases the requirement step of the project is practically missing. Usually a short requirements workshop is proposed to determine the most actors, their goals and initial list use case names. But UP doesn't describe how exactly to capture the information for use case model, based on real customer needs.

Contextual Design, on the other hand, provides a disciplined way that help analysts to find information from the domain. The contextual inquiry, all work models, the interpretation and consolidation sessions are real steps that guide the design team in requirements elicitation. But these steps are not well specified, to show a consistent way of preparing elements of system functionality and communicating them with customers. It is not clearly defined how to include the elements from CD work models into future design of the system.

Regarding these open questions we propose an integrated approach based on main principles of Contextual Design and use case modelling in a new, iterative approach. We use CD in not strictly linear way as it is proposed by [3]. We combine it with elements of use case modelling and apply the iterative approach of the UP.

First, we start with interviews and after the interpretation session we'll have CD models. We'll focus on three of them – Flow Model, Sequence Model and Artifact model. Originally, CD approach continues

with work on the vision, storyboards and UED, but according our observation at this moment the design team hasn't enough knowledge to create the complete vision of the system. This is one of the reasons for the small number of the project that apply CD in its pure shape.

Instead of moving to visioning and storyboarding, as a **second step** of our approach we suggest to continue with more disciplined approach of UP and to use the information from work models for use case modelling. At this step of the project work models are consolidated, showing the common work patterns and strategies across all users—the real work process. As the CD Flow models describe roles and activities they are good start point for defining different actors/roles that operate with the system. The Flow models show all significant users, their activities, the important goals and some artifacts. So we can define the first set of actors for our system.

The activities from the sequence models as well activities from CD Flow models give us the ways that users can use to work with the system at different levels. Sequence models are good source for use case steps; breakdowns can be used for separation of activities or working on alternative scenarios, the intent and the trigger could be successfully incorporated into full use cases description.

As a result the use case model with the most actors, goals, and an initial list of use cases could be named at a brainstorming session. Based on these two models also the boundary of the system could be shaped.

We suggest a template for use cases fully dressed description to be chosen here in order to start with the description of the most important or risky use cases.

At the end of this step the most use cases written in brief format could be ready; 10-20% of the use cases could be written in detail to improve understanding of the scope and complexity.

Applying principles of iterative and incremental development we continue with vision, and storyboards. The concept of *vision* presents in both methods – CD and UP. Here is the place where the two visions – visual from CD and textual from UP could be integrated on a new level as the **third step** of our approach. The combination of images and text description will present the full picture of what the system should do.

Storyboards are other CD elements that use graphical representation. As they could present not only user activities on the system but also some manual practices, initial UI concepts, business rules, etc., it is not easy for the design team to choose which work tasks to show in these series of pictures and text. Now, when we have a set of system use cases with defined preconditions and postconditions

(achieved results) it will be easier to combine them into complete series of activities to achieve a valuable business goal.

The **fourth step** of our approach focuses on the information model of the system. Every system concerns a specific area or specific domain. The real world things and concepts than present the problems, solved by the system should be part of the system model. Domain modeling is this part of the UP that discovers "objects" (classes) that represent those thing and concepts. This is the model that also represents the essential relationship between the concepts.

UP proposes good principles for defining the domain model. The list of conceptual classes provided here, guide the designers to choose the right form of the information structure elements. The description of the use cases is a useful source for the domain elements, too. They could be successfully combined with the elements of several CD work models.

Actually, CD doesn't focus on the information model. But many important object and concepts could be found on flow models. The artifact models also provide good source for many domain elements. The vision could be used for some conceptual classes as well.

The domain model is very important at this step also because it will be the basement for the glossary of the system to ensure the consistent usage of all terms that describe the problem space.

In this way, using the iterative approach at any iteration we consistently describe the static and the dynamic side of the system - the structure and the behavior.

Following the proposed four steps we obtain a complete, correct and detailed description of the system based on the use case and the domain models and we are ready to go to the next level where we can create a detailed design from them.

The steps of the next phase concerns more design activities. A set of UML diagrams will be used to describe the dynamic part of the system - sequence diagrams, activity and state diagrams. The User Environment Design where we can put all different parts of the system and how they supports the user's work is also part of this level. Also prototypes could be used to define the user interface design details.

As this concern design of the system it will be matter of another study.

Results and discussion

In this study, we are trying to define an *analytic strategy that helps designers to* identify customer requirements and specify them in the best possible way. In particular, we integrate several principles of CD with concepts of OO analysis and design and apply them into an iterative

194

process for requirements elicitation based on customer data. We use several design iterations that are necessary for the refinement of the work products from earlier phases of the process.

The presented approach is based on the integration of the concepts, best practices and previous experiences, grounded on a data analytic strategy from OO analysis and design and CD methodology. Several important design activities are followed here:

a) Understanding and specifying the context of use (interviews, interpretation sessions, consolidation sessions, affinity session, etc.);

b) Specifying the user requirements (CD work models – Flow and Sequence, use case description, use case template);

c) Producing design solutions (use case model, domain model, UML diagrams);

d) Evaluating designs against requirements (use case model, domain model).

The presented approach could be used not only as a systems design method but also as an academic research method. We applied it for 6 case studies as a part of the course **Information system analysis and design** for 3-rd year undergraduate students at Faculty of Mathematics and informatics (Sofia University).

Every team started the work on its project with interviews with real customers and followed the steps, defined at the proposed approach. The student conducted interviews with professionals from six different areas to learn about their current practices and to solicit opinions about a potential design of the systems they work on. The interviews and work models preparation followed the principle of the CD, methods and techniques from OO analysis and design were applied, too.

Table 1 shows some numerical results from these projects that confirm the necessity and usefulness of proposed steps.

Table 1: Projects, models and elements

Project	Number of interviews	CD flow models	CD sequence models	Number of actors	Number of use cases (UC)	Number of UC in fully dressed format
Pinterest	2	2	3	6	30	3
Publishing house	3	3	7	12	23	2
Kinder garden	4	4 + 1	4	6	29	3
Meteo System	4	4	3	7	21	2
Internet Auction	2	2	4	3	26	3
Movie system	3	3	2	7	17	1

The benefits of the implementation of our approach on these several projects could be summarized into several main categories:

1. Work in context. All information needed for the projects is embedded into the context. Conducting contextual inquiry at customer workplace the teams examined the context and got the right data.

2. Work in partnership: The interviews use case modelling session, etc. involved users into the analysis and design process and established good relationships. This improved future design decisions and facilitated the acceptance process.

3. Visualization of the concept: The diagrammatic representation of CD work models, use case diagrams, storyboards, graphic vision, etc. helped teams to understand customer data and to communicate them with all stakeholders.

4. Iterative work: As the proposed approach is not strictly linear at every step the team added new details to the system explanation. In this way they supported refinement of the work products from earlier phases of the process.

5. Team work: Working as team during different steps of the project motivated students for better results and improved their soft skills [8], [12].

Conclusion

In this paper we report on a research proposal where several principles of CD are integrated with the concepts of OO analysis and design and applied as iterative approach for requirements elicitation.

The proposed approach can be used for professional and academic projects as well.

In the future we will continue to clarify and improve our approach including more steps and refine the existing ones.

In addition, we can expect an improvement in the efficiency of the method by using tools for describing different models and presenting UML diagrams.

Another step towards improving the effectiveness is to build a mathematical model corresponding to the approach used. Generalized nets (GNs) could be use for modeling here, because they have high applicability in software modeling [1]. The formalism supported by GNs makes it possible to investigate and verify the properties of the constructed models [11].

References
1. Atanassov, K. On Generalized Nets Theory, Prof. Marin Drinov Academic Publishing House, Sofia, 2007
2. Beyer, H., K. Hotzblatt, L. Baker. An Agile User-Centered Method: Rapid Contextual Design, http://www.agileproductdesign.com/useful_papers/beyer_rapid_cd.pdf
3. Beyer, H., K. Holtzblatt. Contextual Design: Defining Customer-Centered Systems, Morgan Kaufmann, 1997
4. Bitter, K., I. Spence. Use Case Modelling, Addison Wesley, 2002
5. Cockburn, A. Writing Effective Use Cases, Addison-Wesley, 2000
6. Holtzblatt, K. Contextual Design: Experience in Real Life, Mensch & Computer, 2001: 1. Fachübergreifende Konferenz. Stuttgart: B.G. Teubner; 2001, pp. 19-22
7. Larman, G., Applying UML and Patterns: An Introduction to Object-Oriented analysis and Design and The Unified Process, 2-nd edition, Prentice Hall PTR, 2002
8. Maneva, N., N. Nikolova. Soft Skills Training for Software People, Proc. of the 7-th Int. conference on Computer Science and Education, July 6-10 2011, pp. 117-129
9. Savoska, S., S. Loskovska. Specific usage of visual data analysis techniques, S3T, Sofija, 2009; pp 234-238
10. Sommerville, I. Software Engineering, 8 ed, Addison Wesley, 2006
11. Todorova, M. Construction of Correct Object-Orientated Programs via Building their Generalized Nets Models, Annual of "Informatics", Section of the Union of Scientists in Bulgaria, vol. 4, 2011, pp. 1-28
12. Todorova, H. Hristov, E. Stefanova, N. Nikolova. How to build up contemporary software professionals (Project-Based Learning in Data Structure and Programming), Second International Conference on Software, Services and Semantic Technologies, September 11-12, 2010, Varna, Bulgaria, pp. 47-54

Author:
Kalinka Mihaylova Kaloyanova
Assoc. Prof. PhD
Sofia University, FMI
359 2 81 61 594
kkaloyanova@fmi.uni-sofia.bg

THE ANALYSIS OF THE DATA OF GEOPHYSICAL RESEARCH OF BOREHOLES BY MEANS OF ARTIFICIAL NEURAL NETWORKS

Y. I. Kuchin, R. I. Muhamedyev, E. L. Muhamedyeva, P. Gricenko, Zh. Nurushev, K. Yakunin

Abstract: *The analysis of geophysical research of boreholes in uranium fields is an important phase of deposits searching, exploration and mining. The method of pre-processing of data and training of artificial neural network are described. This method gives stable results. About 2000 calculation experiments have been made, software and templates for pre-processing of data and results interpretation have been developed. Different methods of normalization, smoothing, excluding shifting data, preparing of training samples were researched. These experiments showed the perspective of neural network approach to solving the problem of rock recognition on stratum-infiltration uranium deposits. Problems of the further experiments which will allow to raise the degree of automation of recognition process and its accuracy are formulated.*

Key words: *geophysical research of boreholes, artificial neural network, pre-processing data, normalization, smoothing.*

Introduction

The method of acidic in situ leaching is widely used for uranium mining in Kazakhstan [1] (fig.1).
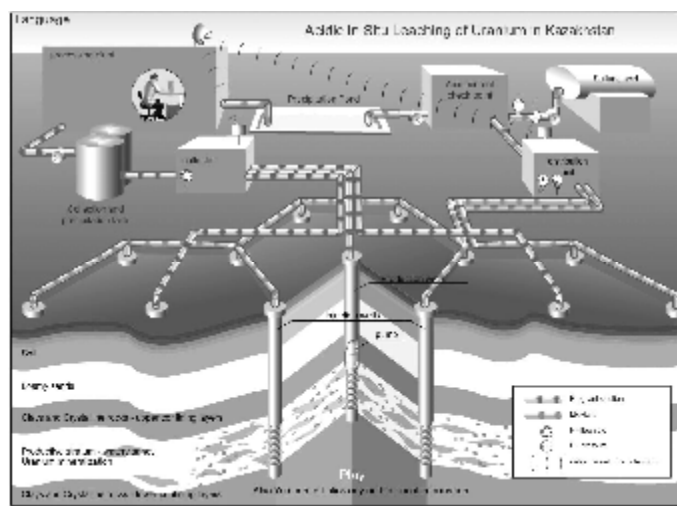
Figure 1. The method of acidic in situ of uranium in Kazakhstan

In this case the efficiency of extraction depends of accuracy and timeliness of data interpretation of geophysical research of boreholes (GRB). The process of logging data interpretation can not be strictly formalized. Computer interpretation methods on the basis of expert estimates are necessary. Among these systems are artificial neural networks (ANN) which have already been used for a wide range of recognition problems [2, 3, 4] including GRB in oil field development [5], and their usage for rock recognition in uranium deposits is offered in [6]. Quality of network essentially depends on its configuration parameters and methods of data processing. The researches results of ANN configuration parameters and pre-processing data methods with using of data from deposit "Budenovsky" are described.

The Basic scheme of Setting of neural net according current task

The neural network setting according current task includes four main stages: **Pre-processing, neural network configuration parameters selection, training of neural network and neural network model estimation** [1]. Pre-processing stage particularly includes normalization, smoothing and errors exception (in our situation it was borehole logs moving and anomalous values deletion).

Network architecture, neuron activation functions and learning algorithm are neural network parameters.

Results depend in considerable extent on chosen method of assessment. In real experiments neural network will work with new data. Thus, the estimation of results on same data set which fully or partially took part in training will not be right. For this reason the percent of correct answers (CCR) on "raw" boreholes (which haven't taken part in training) have been used as metrics of neural network quality.

Pre-processing methods of data

After cleaning of data of abnormal values the next steps of work are the procedures of normalization and centering of data so that every component of initial vector will be within interval from 0 to 1 or -1 to 1.

At a known size of changing of an input variable it is possible to use linear transformation.

$$p = \frac{(x - x_{\min})(b - a)}{(x_{\max} - x_{\min})} + a$$

,

Where [a,b] – the range of acceptable input signals; [x_{min}, x_{max}] – variation range of input variable values; p - converted input signal.

Transformation by sigmoid function or hyperbolic tangent is used in cases of large size of changing of input variables. In some cases can be used modular, positional and functional pre-processing methods [2].

Wavelet data analysis and Fourier transform can be used for cleaning data from the noise.

The technique of experiments

In the process of experiments the impact of following factors was investigated:

- Neural network training algorithm;
- Neural network architecture;
- The type of normalization;
- Smoothing method;
- Special methods of formation of learning samples
- Exception of shifting of logging data

About 2000 calculation experiments have been done. "Floating window of data" with parameters 5+1+5 (there were 5 points above and below the current one plus current point) was an input for network in experiments . Three types of logging data: IL - induction logging, AR - apparent-resistivity logging, NP - method of natural polarization logging were used. The complete size of input vector was 33.

Numerous computing experiments with the application of package Alyuda Neuro Intelligence (http://www.alyuda.com/companyinfo.htm), and with the special software and templates developed by using Python language, have been performed. To implement the wavelet smoothing WaveUtils library was used. Package Alyuda Neuro Intelligence allows to select required network structure, training it by using correct and, in case of need, automatic partition of data into groups (training, validation and testing). Particularly experiments have shown results range on "raw borehole" would be 10% - 15% even for minimum CCR variation on Validation set. Therefore in most cases each experiment was repeated up to 10 times and more in order to get statistically right set of results. The stages of pre-processing of data and formation of learning samples were included in experiments.

Stages of pre-processing of data:

- Deletion of abnormal values
- Nonlinear (X) and linear normalization
- Wavelet analysis (X)
- Removal shift of logging data (X)
- Data formatting into "floating window of data"

Note: stages which either had not influence to results or made it worse are marked with (X).

Stages of training data formation and network training.

1. The combining of boreholes data or using special data set ("ideal borehole")
2. Network architecture selection
3. Algorithm and training parameters selection.

The analysys of Algorithm of neural network training

Some algorithms of neural network training were investigated: Quick Propagation (Quick), Conjugate Gradient Descent(CGD), Quasi-Newton (QN), Limited Memory Quasi-Newton, Levenberg-Marquardt (LM), Online Back Propagation (OBP), Batch Back Propagation. The differences between algorithms are training speed, convergence (some algorithms, for example Leven-M, were not finished - "freeze"), dispersion of results.

As shown in Table 1, algorithm "Conjugate Gradient Descent" on the average gives closer to the best results and at the same time it has minimal dispersion. Therefore that algorithm was chosen as basic for all other experiments.

Table 1. Quality factors of learning algorithm

Parameter	Quick	CGD	QN	LM	OBP
Average	53.9	**58.6**	59.1	"Freeze"	50.8
Dispersion	195.4	**19.7**	42.0	"Freeze"	69.1

Effect of neural network architecture

Neural network architecture is number of neuron in layers, quantity of layers and relations between neurons. The optimal number of neuron in hide layer/layers depends on quantity of input and output neurons. Unfortunately, formal method for determining the optimal number of neurons does not exist. Practically the only method - exhaustive search and checking the various options [2]. However, obtaining an exact solution with any suitable quantity of neurons is impossible in practice, due to the combinatorial explosion in search of solutions. Therefore heuristic methods of package Alyuda were used for practical application. As a result of numerous experiments four-layer neural network architecture were selected. It included an input layer for enter "floating data window", two hidden layers and output layer - [33-39-33-8].

Effect of normalization

One of the most important methods of data pre-processing is normalization. Signal normalization can be linear and nonlinear.

Nonlinear normalization (sigmoid and hyperbolic tangent) had negative effect on quality of recognition (the average percentage of correct recognition decreased by 10%). A linear normalization, in contrast, significantly (15-20%), improved the quality of recognition.

Effect of smoothing

Different methods of smoothing allowing removing of noise and other non informative data content could be used to improve the quality of recognition. There are two discrete transformations often used as methods of smoothing: Fourier analysis and wavelet transform [5].

As this existing selection had almost no high-frequency noise at this stage it was decided not to use the Fourier transform for data smoothing. At the same time, eliminating high-frequency component will lead to informativeness losing of signal (minor fluctuations of rocks will no longer be visible.)

Three (depth of decomposition / scale) variants of wavelet smoothing parameters were used (range of Daubechies wavelet and depth of decomposition / scale).

Table 2. Effect of wavelet smoothing of signal on quality of recognition

Wavelet (range of wavelet – scale)	CCR on set Validation	% on "raw" borehole
1-1	47,4	60,33
1-3	41,8	34,22
7-2	46,2	57,66
No smoothing	48,8	58

The experimental results show us it is no any significant improvements in recognition quality with application of wavelet smoothing. However, wavelet smoothing can be applied in the future, for example, to simplify the identification of AR extremum, which information requires the separation of the input data into categories.

Correction of displacement of logging data

Factor affecting the quality of pre-processing data is shift of IL curve relative to AR and NP curves. Values of AR and NP are fixed one device, and IL another. Usually this displacement is not large and constant (within a single borehole to 0.5 meters up or down). Cross-correlation function was used to determine value of IL displacement relating to AR. As known it has such given form for two finite signals:

$$P(i) = \sum_{j} f(j) * g(i + j)$$

f - the main signal; g – linked signal; i - time shifting of one signal relative to another; P (i) - correlation of shift i.

The most likely is displacement in which signals correlation reaches a maximum. Experiments have shown that the shift affects recognition results only slightly (less than 3%).

Table 3. Analysis of correction of AR data displacement on recognition quality (the part of results is shown)

Boreholes quantity	The number of "raw" borehole	CCR on Validation	CCR on "raw"
8 (n.c.)	9	62%	67%
8 (corr.)	9	59%	62%
8 (n.c.)	7	65%	45%
8 (corr.)	7	60%	43%
8 (n.c.)	7	65%	45%
8 (corr.)	7	68%	48%
8 (n.c.)	6	63%	68%
8 (corr.)	6	63%	68%

Notes:

(corr.) – The data adjusted to the maximum cross-correlation function.

(n.c.) – data with IL curve displaced relative to NP and AR curves (not corrected).

Training was conducted until the learning error (training error) = 90%;

Formation of learning sample

In the interpretation of log data it is very difficult to recognize scarce rocks: dolomite (carbonate rocks, lithotype code 9) and sand-gravel deposits (code 1), whose quantity in the training set is small (less

than half a percent of whole data), which is insufficient for normal network learning .

After preliminary analysis, we decided to create a so-called "an ideal borehole" - a set of data, in which the number of examples of different rocks was the same, and themselves examples were uniformly taken from different boreholes at different depths (within the same horizon). In this case set of learning samples can be reduced significantly.

Note that total we can use 7000 learning samples, which provide an average level of CCR 60%. There were two thousand samples in developed "ideal borehole". Table 5 shows CCR with using some variation of "an ideal borehole", 1+2 - "ideal borehole" where combined first and second variant were used as learning samples (in this case data of scarce rocks were doubled).

Table 4. The part of recognition results with
"an ideal borehole" application

The number of borehole	CCR 1	CCR 2	CCR 1+2
3	48	76	69
4	48	44	62
5	49	53	70
6	58	75	77
7	39	53	65
8	62	69	75
9("raw")	51	67	57

Table 5. Impurity content (in percent %)

The number of borehole	% 3	% 123	% 4	% 12	% 7	% 74	% 47
3	25	31	18	4	9	10	2
4	19	39	5	16	9	6	2
5	30	25	6	19	13	7	1
6	24	25	15	17	4	6	9
7	25	30	20	5	0	10	11
8	26	26	24	10	5	2	6
9("raw")	20	17	13	24	3	9	13

Note. Table columns show the contents of rocks (in percentage) in the corresponding boreholes (rounded to integers). For example column %7 shows the percentage of clay. The table uses the following codes lithotypes: 3 - medium-grained sands, 123 – Sands inequigranular, 4 - from small-grained sands to super fine-grained sands, 12 – various-grained sands with gravel, 7 -clay, 74 - sandy clay, 47 - short-grained clayey sands.

It is evident that using of "ideal borehole" can achieve approximately the same level of recognition (60%) with three times smaller data amount.

Conclusion

The set of experiments allows us to recommend the developed methodology for future experiments on the setting of neural networks for problems of uranium boreholes logging. But we have to take into account the following limitations.

Due to the random initialization of weights of a neural network the difference of results are less than 5% can not be considered sufficient reason for choosing a particular method of data pre-processing due to the significant influence of random initial distribution of weights.

When we place data to network input as a floating window with parameters 5+1+5, small shift (up to 10 points) of IL curve relative to AR and NP curves had not affect the recognition results, because the network is "trained" to detect such shifts.

There are not significant improvements in recognition quality on application of wavelet smoothing.

The effective method of preparation of samples is formation of special learning sample ("an ideal borehole").

For achievement of the best result of recognition it is necessary to solve the following problems:

Problem 1. Estimate limits of accuracy of recognition of the data.

Problem 2. Study possibility of use of the general geological data to improve the quality of interpretation.

Problem 3. Develop methods and algorithms to eliminate anomalies in the initial data.

Problem 4. Develop methods of data pre-processing on the "acidified" boreholes, methods of predicting spreading of leach solutions, etc.

References

1. Yashin, S. A. Underground holes uranium leaching at the deposits of the Kazakhstan. "Gorny Zurnal" Scientific-technical and industrial journal.- 2008. N 3, ISSN 0017-2278. p. 45-49
2. Neurocomputers. Tutorial. M.: MSTU named after Bauman, 2004.-320 p.
3. Stuart Russell and Peter Norvig. Artificial Intelligence: A modern approach. Pearson Edition, Inc., Upper Saddle River, New Jersey 07458. 2010. ISBN-10: 0136042597, ISBN-13: 9780136042594
4. M. Tim Jones. Artificial Intelligence: A Systems Approach. INFINITY SCIENCE PRESS LLC Hingham, Massachusetts, New Delhi, 2008. ISBN: 978-0-9778582-3-1
5. Kostikov, D. V. Tools of interpretation of geophysical researches of boreholes on the basis of the transformed logging diagrams by means of a multilayered neural network.-M.: Russian state library, 2007.- 189 p.
6. Kuchin, Y., Muhamedyev, R., Muhamedyeva, L. Interpretation of log data of boreholes. The 9th INTERNATIONAL CONFERENCE: INFORMATION TECHNOLOGIES AND MANAGEMENT 2011, April 14-15, 2011, Information Systems Management Institute, Riga, Latvia, ISSN 1

Authors:

Y. I. Kuchin, senior engineer "GeoTechnoServiss", Almaty

R. I. Muhamedyev, professor, head of chair, IT department, International IT University, Almaty

E. L. Muhamedyeva, m.c.s., senior researcher, International IT University, Almaty

P. Gricenko, student-researcher, International IT University, Almaty

Zh. Nurushev, student-researcher, International IT University, Almaty

K. Yakunin, student-researcher, International IT University, Almaty

Contacts:

Ravil Muhamedyev, email: ravil.muhamedyev@gmail.com

phones: +7 701 3393720, +7 777 2399788

INFORMATICS AND SOFTWARE BUSINESS –
A MUTUALLY BENEFICIAL COLLABORATION

N. Maneva, Kr. Manev

Abstract: The paper describes how the representatives from two communities – computer science researchers and software developers, can work together, exchanging approaches, techniques and tools so as to enrich their knowledge and skills and to improve the results of their professional performance. An example for establishment of mutually-beneficial collaboration, involving participants from a research institute and a successful software company has been given.

Keywords: Informatics, Formal methods in software engineering, business best practices.

1. Introduction

The role of computers in our everyday life and work is tremendous. That is why the concern of the society about the quality of used computer systems should be taken into account by all professionals, responsible for design, development, distribution and maintenance of such systems and services, related with them.

This paper reveals some possible relationships and interdependencies among representatives of two areas – computer science and business. In Section 2 the influence of informatics on software business is considered. In Section 3 some pragmatic advices, identified in business, but validated as helpful and applicable for scientific research, too, are described. Section 4 summarizes the results of a systematically planned and accomplished collaboration between a research institute (Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences) and a leading Bulgarian software company. In the Conclusion a few ideas for further development of joint projects between academia and business are shared.

2. How informatics can help software business

The expanding use of computer systems requires high quality software, because the main functionalities of such systems are achieved via software components rather than hardware implementations [4]. In order to meet these requirements, the contemporary softwaresystems should be developed in a systematic way, involving methods and tools, recognized as effective. Unfortunately, many individuals and companies still develop software haphazardly, ignoring the power of the modern

methods, even in projects, building systems to service today's most advanced technologies [6]. As stated in [4], themaincontributionsof computer science researchers till now are in programming languages, algorithm complexity theory,modellingand model checking,program proving, concurrency, real-time systems,and distributedcomputing. Some remarkable results are obtained in case of development of special-purpose softwareas compilers,database management systems, virus checkers, etc.

2.1. Formal methods in SE - state of the art

The most significant research results are connected with *formal methods* in Software Engineering - the use of mathematical models for analysis and verification duringthe consecutivestages of the software life-cycle.This approach, recognized as a powerful, comprises a specification notation with formal semantics, along with a deductive apparatus for reasoning and is used to specify, design, analyze and ultimately implement a hardware, software (or a hybrid) system [4].The state of the art in the industrial use of formal methods is presented in [7]. Here we only givesome conclusionsmade after analysis of survey data aboutthe impact offormal methods on improvements in practice:

• Among the most preferable application domains, in which the formal method have been successfully applied, are transport, financial sector, defence, telecommunications, office and administration. 20% of responses in the survey additionally indicated thatthe projects, involving formal methods,are related to software development tools themselves, such as operating systems,compilers, and CASE tools. Further 10% are related to computing applicationswithin the domain, such as high-performance computing, runtime code optimisation,file system replication, access control, communications protocols, and microcomputerdesign.

• Two main reasons for the restricted use of formal methods have been identified: insufficient theoretical (mathematical and algorithmic) qualification of software developers and lack of supported tools (stand-alone or integrated).

2.2. Our approach to usage of formal methods in SE

We share the understanding that it is necessary to create awarenessof all people, involved in software development, that the power of theoretical knowledge (both mathematical and algorithmic) is high and should be used to find a solution of any hard problem, arising during software life cycle. This objective should be defined as a didactic one for computer science students to be achieved during their study at universities.

Togetherwith pure investigation work on inventing new methods, the additional responsibilities of the researchers in the field of computer science should be:

- Using appropriate communication channels, to disseminate all valuable scientific results among both computer science students and software practitioners so as to keep them informed about novelties in the area;

- For some already existing and validated (theoretically or empirically) methods, to develop a procedure for its transfer to practice, and in case it is possible, tocreatesome tools, facilitatingthe methods application.

Following this recommendations, we develop a "3 in 1" approach, comprising a *formal method*, a *systematic step-wise procedure* for its accomplishment, facilitated by a *set of tools*. The methodology is called INSPIRE (an abbreviation of the first letters of the names of its characteristics – **I**ncremental, **N**eat, **S**calable, **P**ermanent, **I**ntegrated, **R**ight, **E**stimable).The essence of the methodology, presented in [5], is that for a selected SE activityit usesin a continuously expanded way a well-defined and flexible formal method. This methodcan be combined with some other activity-specific methods so as to achieve a right (validated) and measurable technological process.

Next follows a brief description of that formal method, built into the methodology – the Comparative Analysis (CA) method. Generally speaking, Itsupports the reasonable choice and shares the objectives and techniques of the multi-criteria decision making.

Formally, t*he Comparative Analysis* (CA) is a study of the quality content of a set of homogeneous objects and their mutual comparison so as to select the best, to rank them or to classify each object to one of the predefined quality categories.

The compared objects can be products, processes or resources, identified as significant for the SE activity under consideration.

In order to apply the CA method, we distinguish two main roles: the ***Analyst***, responsible for all aspects of CA implementation, and a ***CA customer*** - a single person or a group of persons, who should make a decision in a given situation. For each of these roles the required qualification, competencies and psychological traits are defined, together with corresponding typical tasks and responsibilities.

The context of the desired comparative analysis is specified through a *case*, described by the following 6 elements:

case = { View, Goal, Object, Competitors, Task, Level}

The ***View*** describes the customer's role and the perspective from which the comparative analysis will be performed, the ***Goal*** reveals the

purpose to perform the CA, the **Object** presents the studied item. The **Competitors** are instances of the object, compared through algorithms with different complexity so as to perform the **Task** – to find the best, to rank or classify the Competitors in a way, specified by the **Level**.

In order to illustrate the CA method, we will give two examples of its use within the project „Automated business rules and processes extractionfrom software source code". First, the CA has been applied for constructin a quality model for the object „computer system". Following[3], the created model comprises four quality factors: functionality, performance, cost,and dependability.During the second experiment, the Comparative analysis has been used for constructing a quality model for the object "method for business rules extraction". The obtained hierarchical model comprises three quality factors (applicability, validity, utility), further decomposed to a number of characteristics, specific for the object "method". A number of CA cases have been defined so as to rank the competitive methods, evaluating them from three different points of view – those of Policy Maker, Business Analyst and Rules Extractor. Summarizing the results of performed CA, the method of static analysis has been selected as the most appropriate. Our experiments confirmed the expert's recognition of its value [4] as an effective formal method, which is appropriate (solving the task), sound (the results obtained can be trusted), nonintrusive (end usersdo not have to alter their programmingmethods), feasible (applicablein any industrial environment) and scalable (can be used to millions of lines of code asfound in existinglegacysoftware).

The methodologyINSPIRE has been successfully applied for a number of different in scope and complexity Software Engineering activities - software quality assurance, usability assurance, outsourcing development, models evaluation, Informatics teaching,etc., thus proving its feasibility.

3. How business can help the research in Informatics

The collaboration between research and business organizations can be considered as a beneficial for academia in two directions: managerial and methodological.

From business point of view, there is an idea to change the managerial framework, introducing the so called**project-based approach to scientific research.** Instead of traditional researchwithout rigid financial and time constraints, many research activities can be accomplished through projects with clearly stated goals, budget and duration, pre-defined by the funding sources. Such project-driven

research will change entirely the style of scientific work, especially in the field of Computer science, where the pressure for innovative and transferable to practice results is high. In order to provide the desired project's deliverables in time, the scientists, involved in such projects, should enrich their professional competencies with a piece of management knowledge and special "soft skills" (self-management, problem solving, decision making, team-work), which will assure the successful work on the project.

The methodological aspect of collaboration is the following: A number of pragmatic rules and best practices, proved to be successful in business area, can influence and shape in a right way the scientific research process in the field of Informatics. Such useful business practices to be applied in aninnovative project-based scientific research,as recommended in [2], can be:

- Define a set of topics to be studied, evaluate them from different points of view and select only some of them to proceed with;
- For selected topics, state a few clearly defined indicators for success in research;
- Apply an incremental approach - step, don't leap;
- Stick to the principle "Invest a little, learn a lot". It requires a rigorous decision making to judge whether to continue or to cancel the project and how to control the investments in accordance with effectiveness of the work;
- Consider the human factor (peopleware) as a crucial success factor in any research activity.

4. A Case Study: some lessons learned

We are going to present an example for establishment of mutually-beneficial collaboration, involving some enthusiastic participants froma research institute (Institute of mathematics and informatics, BAS) and a successful software company, connected to many companies with IT–intensive business, namely MusalaSoft OOD.This collaboration has been accomplished in two directions – joint scientific projects and educational programs.

Two*joint projects*have been implemented: an industrial project № 5ИФ-02-3, supported by the National Innovative Fund, attached to the Bulgarian Ministry of Economy and Energy and the project „Automated business rules and processes extractionfrom software source code", which is supported by the National Scientific Research Fund of Bulgaria under the contract ДТК 02-69/2009. In these projectseach partner has been responsible for work, requiring the respective competences. The

scientists from IMI provide professional knowledge, expertize and consultations for scientific approaches and methods to be used, and software developers from Musala Soft perform the tools prototyping and experiments to examine the applicability of the methods within a real-life environment. The results of the partnership till now are encouraging – the first project has been successfully finalized, the second project is at the end of the Phase 1. A number of papers, describing the results of the co-work of participants, have been accepted for publication.

The collaboration in educational activities has been organized in a systematic way, which is described in detail in [1]. Among the most significant results at doctoral, Master's and Bachelor's degree are the following:

- IMI assures the tutoring of 2 PhD students - company's employees, under contracts for encouraging the scientific development of young people;
- Company's representatives proposed some topics, which can be further developed as Master's Thesis works within the program "IT Management" and then participated as Jury's members for course-works and Master's thesis evaluation;
- Some successful managers from the Marketing and Customer Support Departments of the company have been invited lecturers for a few SE courses. Their lectures cover some modern approachesin software develipment;
- Selected company's employees participated in defining the assignments for teams of students, working on projects for the practical seminars;
- Discussion of some real-life business-oriented topics so as to be included in SE course syllabus.

In order to create awareness about the crucial role of science and innovations and to encourage collaboration with other partners, the IMI organizes a conference "Mathematics in Industry". The Institute is a co-organizerof a conference "Informatics in the scientific knowledge". During these conferences many ideas for joint research and development have been exchanged and the most appropriate forms of partnerships have been discussed.

5. Conclusion

The paper shows how researchers in Informatics and software developers can work together so as to assure quality software. An example of mutually beneficial collaboration between a research Institute

and a software company has been given, showing that such collaboration is feasible and fruitful.

Some possible directions for future work are:

• To continue the IMI policy of collaborative work, involving other IT companies and professional organizations, like BASCOM and BYTE;

• To start or join for participation in some projects with high social impact, e.g. e-Government projects;

• To plan and accomplish a systematic dissemination of the significant scientific results, by different communication channels, including Internet sites and social networks.

6. Acknowledgments

References
1. Манева, Н., Кр. Манев. Тристранно партньорство в обучението по Софтуерни технологии–предизвикателства и резултати. *Сб. доклади на нац.конференция „Образованието в информационното общество"*, Пловдив, 31 май – 1 юни, 2012, с. 205–215.
2. Anthony, Sc. et al. Mapping Your Innovation Strategy. In *Harvard Business review on Business Model Innovation*, Harvard Business School Publishing Corporation, 2010.
3. Avizienis, A., J.-C. Laprie and B. Randell: *Fundamental Concepts of Dependability*. Research Report No 1145, LAAS-CNRS, April 2001.
4. Hinchey M. et al.Software Engineering and Formal Methods.*Communications of ACM*, September 2008, Vol.51, no. 9, pp. 54–59.
5. Maneva, N. INSPIRE: A Software Engineering Methodology. *Journal of Information Technologies and Control*, № 1, 2005, pp. 22–26.
6. Pressman, R. *Software engineering – A Practitioner's Approach*, Seventh edition, McGraw Hill, 2010.
7. Woodcock, J. et al. Formal Methods: Practice and Experience. *ACM Computing Surveys*, Vol. 41, No. 4, Article 19, October 2009.

Assoc. Prof. PhD Neli Miltcheva Maneva,
Institute of Mathematics and Informatics, BAS,
e-mail: neman@math.bas.bg

Assoc. Prof. PhD
KrassimirNedeltchevManev,
Faculty of Mathematics and Informatics, Sofia University,
e-mail: manev@fmi.uni-sofia.bg

NANOELECTRONICS AND EFFICIENCY OF INFORMATION SYSTEMS OF MANAGEMENT

Alexander Mrochko

Abstract: *Information systems (IS) are the basis of effective management of complex dynamic systems of any type (technical, economic, social, and biological, etc.). The structure of IS includes the following subsystems: reception and transmission of information, processing and transformation of information, storage of information. The basic technical element of the above-mentioned subsystems is the electronic device. Since the emergence of the first electronic devices, the research thought has been aimed at the improvement of their application properties: reduction of sizes, increase of the performance rate and reduction of the consumption power.*

Key words: *nanoelectronics, electronic devices, information systems, efficiency indicators*

The process of forming of global information space has been developing along with the development of socio-economic and scientific-technological development of international community. Among the main statements of Okinawa Charter on the Global Information Society adopted by the leaders of the "Big Eight" there is the following: any individual, a group of individuals, a firm or an organization in any point of a country and at any time can receive any information and knowledge, which are necessary for their life and solution of personal and socially significant tasks, based on the automated access and communication systems, for a fee or free of charge. The increasing technical capabilities play significant role in the formation of the global information space. These capabilities define the process of encoding-decoding, formalization, accumulation, storage and transmission of information, as well as its accessibility and relative cheapness of more sophisticated technical tools incorporated in information system.

Information systems (IS) are the basis of effective management of complex dynamic systems of any type (technical, economic, social, and biological, etc.). The notion "information systems" can be defined as a set of interrelated components which collect, process, store and distribute information.

Characteristics of the quality of information usually are a list of properties of information:
- representativeness;
- meaningfulness;
- completeness;
- availability;

214

- topicality;
- timeliness;
- accuracy;
- reliability;
- stability;
- value;
- etc.

Information is the product of IS, thus, all the aforementioned properties of information are closely related to the indexes of information systems efficiency. The generalizing indicators of the information system's efficiency are the indicators of the economic efficiency. Calculation of expenses does not present any difficulty but calculations of results remains a complex problem.

Usually, for evaluating the efficiency of the IS, there are used some particular indicators of efficiency which characterize:

- *Pragmatic efficiency* – indicators of reliability of information; indicators of the information system safety; indicators of the accuracy of calculations and of the information transformation; indicators of the completeness of the information formation; performance indicators.

- *Technical efficiency* – reliability indicators, functionality, number of the served subscribers, productivity, capacity, clock frequency, temporary delays, memory capacity, operational characteristics, technologies of service, etc.

- *Technological efficiency* – simplicity and adaptability to the development and creation of the system.

- *Operational efficiency* – convenience of the system use and maintenance.

Technically **Information System** includes all the devices providing obtaining, processing, storage and transmission of information:

- various sensors;
- systems of conversion and processing based on computer technology;
- storage systems;
- means of radio and telecommunications.

Modern information systems use electric signals (analog or digital) to represent information. The electronic devices, which principle of action is based on the interaction of charged particles with electromagnetic fields, make the technical basis of information systems. The area of science and technology on the creation of electronic devices (vacuum

and solid state devices) is called electronics, and it emerged in the early XX century.

There are two types of electronic components that we come across namely Passive (Resistors, Capacitors, Inductors, etc.) and Active (Diodes, Transistors, etc.) components. These components can be combined in different configurations by interconnecting them with conducting wires to build different useful Electronic devices and circuits (rectifiers, amplifiers, oscillators, filtrs etc.).

The first electronic device, a vacuum diode (Fig.1), was patented in 1904 by D.A.Fleming
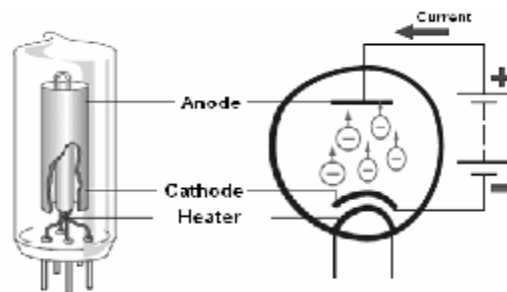
Fig. 1. Diode valve construction and symbolic representation

and original three-element device, a vacuum triode (Fig.2), was patented in 1906 by Lee De Forest.
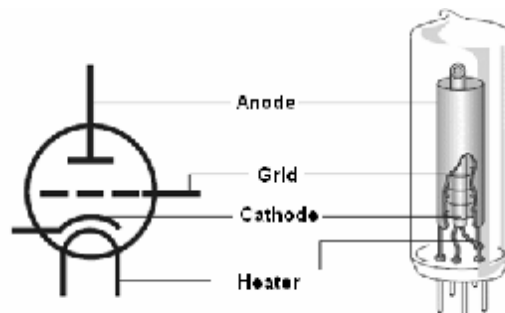
Fig. 2. Triode valve construction and symbolic representation

Fig. 3. Vacuum tube (diode and triode)

In 1946 the World's First Electronic Computer - ENIAC (Electronical Numerical Integrator and Calculator) , weighing 27 tons, dimensions 2.6 m × 0.9 m × 26 m (taking up 63 m^2) was created. Its components included 17,468 vacuum tubes, 7,200 crystal diodes, 1,500 relays, 70,000 resistors, 10,000 capacitors and approximately 5,000,000 hand-soldered joints (Fig.4).

Fig. 4. ENIAC, 1946

Since the emergence of first electronic devices, the research has been aimed at the improvement of their application properties: reduction of sizes, increase of the performance rate and reduction of the consumption power.

The invention and practical application of the first semiconductor device – transistor, which was demonstrated at Bell Telephone Laboratories in 1948, fostered the intensive development of solid-state electronics, especially semiconductor one.

The emergence of the new direction in electronics, which is known as microelectronics, is connected with the first integrated circuits (IC) on silicon in the late 1950s. In recent decades, microelectronics has reached enormous success in the terms of high density and miniaturization of elements (Fig.5).

Fig. 5. Various components of microelectronics

In fact, the well-known Moor's law, doubling of the number of transistors that can be packed in integrated circuits every 18 months or so, has remained fair during the past 40 years and is still true today (Fig.6) [1].
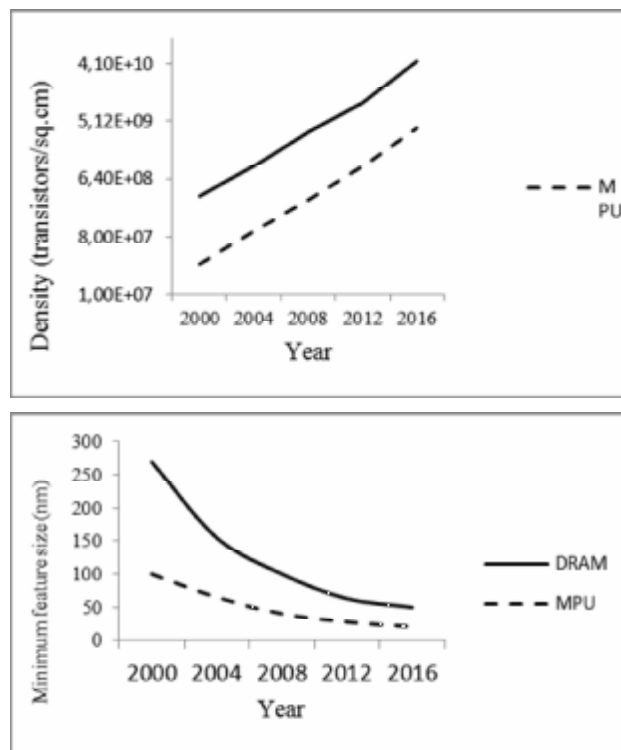
Fig. 6. Trends in density and size of electronic devices

At present, the sizes of electronic components in integrated chips have already attained the 50-nm range and the density of one billion transistors per square centimeter has been reached [1,2].

Due to the achievements in microelectronics, it has become possible to develop sophisticated electronic systems with a high degree of integration, high rate of data processing, low energy consumption and, therefore, highly efficient information systems on the whole.

However, further reduction of the sizes of transistors meets certain difficulties. The main challenges lie in the area of the gate dielectric, the gate electrodes, the substrate and device structure, and the devices interconnections. Also, as MOSFETs are shrunk, they become sensitive to the fine structure of random distribution of dopants in the devices.

The need for smaller and faster electronic devices has given life in the recent years to the new branch of nanoelectronics. Researchers and developers are interested in nanoelectronic properties of materials for the purpose of communication, computation, storage or control.

In the recent years there have been developed electronic devices (single-electron device, resonant tunneling diodes and different spintronic devices) which are undergoing laboratory research. Due to their characteristics, these devices present a serious competition to the main elements of information systems - to traditional semi-conductor transistors. In the nearest future there is expected rapid development of Carbon nanoelectronics and emergence of the key electronic components on the basis of carbon nanotubes.

In August 2007 the group of researchers from University of California at Berkeley have constructed a fully functional, fully integrated radio receiver from a single carbon nanotube [3]. The nanotube serves simultaneously as all essential components of a radio: antenna, tunable band-pass filter, amplifier, and demodulator. Using carrier waves in the commercially relevant 40-400 MHz range and both frequency and amplitude modulation techniques, we demonstrate successful music and voice reception.
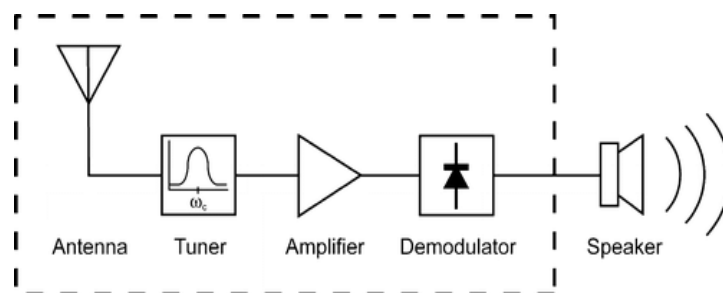
Fig. 7. All in one nanotube radio

Thus, the emergence and rapid development of nanoelectronics provides new opportunities for the improvement of some particular indicators of Technical, Technological and Operational efficiency of Information Systems, reduction of sizes, increase of the performance rate and reduction of the consumption power of the electronic components.

References
1. Daniel Minoli. Nanotechnology applications to telecommunications and networking. John Wiley & Sons, Inc., 2006. 487 p.
2. Борисенко, В. Е. Наноэлектроника. – М.: БИНОМ. 2009, 223 с.
3. K. Jensen, J. Weldon, H. Garcia, and A. Zettl. Nanotube radio. Department of Physics, Center of Integrated Nanomechanical Systems, University of California at Berkeley. August 21, 2007.

Author:
Alexander Mrochko,
Dr.Sc.ing , Ass.prof.
Information Systems Management Institute
Address: Ludzas 91, LV-1003, Riga, Latvia,
Tel.: (371) 29614403.   E-mail: _aleksandrs.mrocko@isma.lv_

PRELIMINARY RESULTS OF THE DEVELOPMENT OF PORTAL "ACTIVE LONGEVITY OF KAZAKHSTAN POPULATION"

**R. I. Muhamedyev , V. G. Dmitriyev, A. T. Mansharipova,
M. M. Maratov, E. L. Muhamedyeva, D. A. Shynybekov,
I. M. Ualiyeva, D. V. Zagulova**

Abstract: *Increasing the duration of active life is one of the most important problems in modern gerontology in the world, including Kazakhstan [1]. First, the solution of these problems is associated with using information and communication technologies (ICT) [2-5]. Gerontological RK portal, which has no analogues in countries of the former USSR, will include information on major aspects of the aging in RK. Portal will be a platform for exchange of experiences, publishing of research results, for getting generalized and personalized information. This paper describes the main elements of the portal and presented preliminary results of frequency occurrence analysis of cardiovascular diseases in elderly age and income level of the respondent on the basis of epidemiological screening, performed on the basis of data collected on our site.*

Key words: *gerontology, Kazakhstan population, information technology, cardiovascular diseases.*

Work actuality

Gerontological research in the world are associated with the rapid aging of the population in developed countries. This process is caused not only by decreasing the birth rate, but also by an increasing in life expectancy. For example, according to European experts in 2060 one third of Europe's population will reach age 80 and elder. At the same time there are great part of single people and people who need intensive medical care in a given layer of the large population. In this regard, the main tasks set forth by the scientific and medical community developed countries are preventing the social isolation of senior citizens, creating conditions for maintaining an active lifestyle for elder people and their participation in the labor market.

To achieve these ambitious goals the world community make a range of initiatives by medical, social and technological types of the widespread using of modern information and communication technologies. It is not just proper measures to maintain health, but also the initiatives in the field of business, for employment, using of ICT to reduce the costs of health care and care for elderly patients. As a result of these efforts during the period from 2010 to 2013 only in Europe, public and private investment for research and innovation to improve lives of elder people will be more than 1 billion euros.

Kazakhstan is a developing country, but country's population is also aging. In Kazakhstan, as of January 1, 2010 the number of people aged 65 years and older was 7.14% of the total population of the country. UN experts consider Kazakhstan is state with accelerated aging, and 25% of elder people are expected in the country according to forecasts for 2050.

Because of this, the problem of prolongation of active longevity, reducing the costs of health care for elder people, increasing their demand in the labor market in old age also stay in front of Kazakh society.

Improving care for elder people, comprehensive solution of their medical-biological, social, and psychological aspects is one of the priorities defined by the State program "Salamatty Kazakhstan" for 2011-2015, approved by Presidential Decree of Kazakhstan for 29.11.2010, № 1113.

Economic analysis shows that we need to find ways and new technologies, which with improving the quality of services would be able to reduce costs. Initiatives that are offered in different regions are usually connected with using of information and communication technologies (ICT).

Although the group of elder persons is heterogeneous by education, income, and type of illnesses associated with age, older people are the group of maximum risk in the sense of exclusion from the benefits of the Information Society. Recent study, for example, found that more than 60% of people over 50 in Europe feel that their needs are not taken into account by current ICT services and equipment [4].

Policies and initiatives aimed at promoting conditions and technology for elder people in the information society can achieve a triple purpose:

- facilitating of easy using of the information society tools and services for elderly people, among other things removes barriers, making ICT tools easier for everyone and encourages people to look for better, increasing their independence, improving health, increasing the ability to participate in work and be active in society;

- ICT can help to increase the efficiency and quality of social services and health in aging society, and promote financial sustainability of these services in the future;

- because of almost global phenomenon of aging new ICT products and services to ensure the needs of aging society provide an opportunity in corresponding industries to become leaders and exporters in the global market.

At the same time we have to point that the potential of ICT in gerontological market is very high, may say this is huge: the fact is Europeans aged 65+ own funds over then €3000 milliard. In Kazakhstan, the situation is different, but due to the increasing of elderly population (1634974 people at the beginning of 2011), an increasing of Internet usage, income growing, as well as increasing number of specialists in the field of gerontology and geriatrics, the potential of ICT will also be high.

Thus, increasing duration of people active life is one of the most important problems of modern gerontology in the world, including Kazakhstan.

Solution of these problems is connected with application of information-communication technologies (ICT). The work of geronto-logical portal development is in progress. This portal doesn't have any analog in the countries the former USSR and in neighboring countries.

Objectives and functions of the portal

The portal is supposed to be as an input point to "the gerontological space" of Kazakhstan, including information about major aspects of aging of Kazakhstan population. It will be as a platform for an exchange of experience, the publication of research results, for generalized and personified information reception about elder person health.

Based on the collected information using methods of scientific forecasting, searching patterns in the data and forecasting in processes of aging by experts in gerontology and geriatrics, expert systems on various aspects of aging process will be developed.

Portal users at this stage:
- Doctors (managers for interviewers)
- Interviewers (persons who gather and enter questionnaires)
- Managers

Portal functions at this stage:
1. Questionnaires selection according to various criteria
2. Downloading forms data in .xls format
3. Questionnaires input, editing, deleting
4. Get information about interviewers
5. Get statistics from questionnaires entered into data base

Prospective users of the portal in the future (additional to initial portal users):
- Specialists in the gerontology

- Any other medical specialists, practitioners, Interns, residents, graduate students
- Health facilities
- Sanatorium and rehabilitation centers
- Social services
- Elder citizen, their relatives and carers
- Organizations that use volunteer work
- Any interested organizations and individuals, including entrepreneurs, merchants, sponsors and patrons, etc.
- Health insurance
- Medical equipment and medicines manufacturers and distributors.

Possible functions of portal in future:
- Ability to self-entry questionnaire by patients or their relatives
- Statistical information on the questionnaires from portal database
- Ability to analytics
- Self-diagnosis
- Placement of scientific publications and scientific journal of the medical community
- Relation: patient - portal - clinic. Information exchange between the patient and the portal, the patient and the hospital, the clinic and the portal
- Portal users forum on gerontology or other medical issues

Questionnaire data gathering and store

Personal data about health of elderly and senior patients in Kazakhstan collecting and store is a part of the portal. That data required for the further analysis and prediction of the aging process on Kazakhstan territory - information system " questionnaire".

Figure 1. presents the local system scheme of questionnaires gathering.
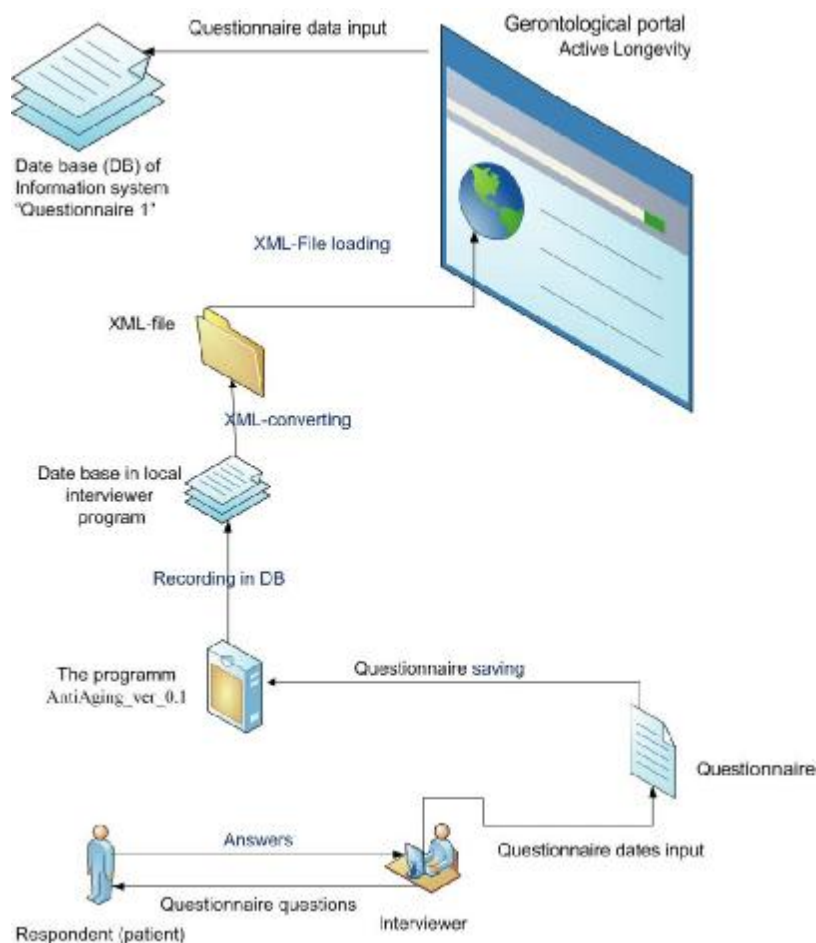
Fig. 1. Process of questionnaire data about health
of elderly and senior citizens collecting and storage

The information system provides a systematic collection of personal data of patients with subsequent loading into the database.

Preliminary results

The results obtained during collecting and entering data allow us to get some results at this stage already.

The priorities in health and social safety at the present stage are health promotion, prevention of disease and disability, the development of rehabilitation. Public health is the most important indicator of society

well-being and depends on many factors: the environment, labor conditions, level of satisfaction of population needs in food, clothing and other goods, social goods, leisure, etc. Therefore, to solve health problems is not enough only budgetary funding model of medicine. Efforts in this direction may be weaken by unsufficient contribution to the health by the population. The last may be associated with one hand, the reluctance to healthy lifestyles, and on the other hand, the lack of such opportunities in the population, in particular material.

Cardiovascular diseases (CVD) is considered as one of the major problems in modern world. Therefore the income influence on frequency of Cardiovascular diseases was investigated in different age groups.

Income level was evaluate with questionnaire where respondent by himself defined his level of income choosing an option from the list:

"Money is hardly enough for food"

"It is enough money for food but for clothing isn't."

"It is enough money for food, clothing and small appliances "

"It is enough money for large household appliances "

"It is enough money for all except estate"

There are no any problems with money

Mathematical processing of the results was performed using the software package «STATISTICA 8.0»

Analysis of the interactions showed that really the income of the population and frequency of CVD in groups of men and women has negative correlative relationship. However, at the same time, it was found that the level of income in the studied group has an inverse correlation relationship with age. (Table 1a, 1b).

Table 1a. Relation between age, income and CVD
(Spearman rank correlation coefficients) in the group of men

	CVD	Income
Age	0,24 (p=7,8E-10)	-0,18 (p=0,004)
CVD		-0,33 (p=2,7E-05)

Table 1a. Relation between age, income and CVD
(Spearman rank correlation coefficients) in the group of women (p<0,05)

	CVD	Income
Age	0,26 (p=1,1E-19)	-0,17 (p=0,002)
CVD		-0,28 (p=4,7E-08)

Therefore, to eliminate the influence of age factor on the relationship of income and CVD frequency was examined in different

age groups. For this age group formed in such a way that, firstly, the age in subgroups with cardiovascular diseases and without them did not differ, and, secondly, that in selected age subgroups was not statistically significant correlation of age and income.

Thus were formed groups in which the incidence of cardiovascular disease and the level of income were not associated with age. The group "Men" was divided into two groups: one group consisted of men from 45 to 55, the other elder then 64 years (Table 2a, 2b)

Table 2a. The result of comparisons of income levels and age in groups of 45-56 years men with CVD (group1) and without CVD (group2)

	Ranges sum of group1	Ranges sum of group2	p-level	Valid N group1	Valid N group2	Median group1	Median group2
Income	104,0	1436,0	0,004	8	47	2,5	5
Age	696,0	2154,0	0,10	15	60	50,5	49

Table 2b. The result of comparisons of income levels and age in groups of men over 65 years with CVD (group1) and without CVD (group2)

	Ranges sum of group1	Ranges sum of group2	p-level	Valid N group1	Valid N group2	Median group1	Median group2
Age	3097,000	3119,000	0,4	50	49	73	72
Income	2929,000	1922,000	0,01	44	46	2	3

Among women only one subgroup followed specified conditions. It included women elder than 55. In the subgroup of women younger than 55 years were no any statistically significant relationship between the frequency CVD occurrence, age and income.

Table 3. The result of comparisons of income levels and age in groups of women elder than 55 years with CVD (group1) and without CVD (group2).

	Ranges sum of group1	Ranges sum of group2	p-level	Valid N group1	Valid N group2	Median group1	Median group2
Age	17195,50	23274,50	0,63	123	161	70	67
Income	14629,50	12166,50	0,00003	108	123	2	3

Note: Rank criterion of Mann-Whitney was used as a criterion for identifying differences in characteristics levels in sub groups.

Analysis of CVD frequency at different income levels using two-tailed Fisher's Exact Test revealed statistically significant association between these parameters in groups of men of 45-55 (p = 0,002) and over 65 years (p = 0.04). (Table 3a, 3b). In the analysis due to lack of data after checking for homogeneity were combined the data about income, which was higher than "money is hardly enough for food".

Table 3a. Frequency of occurrence of cardiovascular diseases for men eleder than 45-55 in groups with different income levels

Income level	There are CVD	There are not CVD
"Money is hardly enough for food" (1 unit.)	4	1
The level is higher then "Money is hardly enough for food"	4	36

Table 3b. Frequency of diseases occurrence for men elder than 65 years in groups with different income levels

Income level	There are CVD	There are not CVD
"Money is hardly enough for food" (1 unit.)	14	5
The level is higher then "Money is hardly enough for food"	35	42

In the group of women elder than 55 was found an association between the CVD frequency and income using the criterion χ^2 ($\chi^2 = 23,3$, p = 0,007) (Table 3a, 3b).

Table 3a. Frequency of disease occurrence for women elder than 55 in groups with different income levels

Income level	There are CVD	There are not CVD
Money is hardly enough for food	36,6 % (45)	10,2% (11)
It is enough money for food and for clothing isn't	24,4% (30)	26,9% (29)
It is enough money for food, clothing and small appliances	19,5% (24)	30,6% (33)
It is enough money for large household appliances + It is enough money for all except estate	3,2% (4)	5,6% (6)
There are no any problems with money	16,3% (20)	26,9% (29)
Total in group	100% (123)	100% (108)

Table 3a. Frequency of diseases occurrence for women elder
than 55 in groups with different income levels: expected numbers

Income level	**There are** CVD	**There are** **not** CVD
Money is hardly enough for food	24,2% (29,82)	24,2% (26,18)
It is enough money for food and for clothing isn't	25,5% (31,42)	25,5% (27,58)
It is enough money for food, clothing and small appliances	24,7% (30,35)	24,7% (26,65)
It is enough money for large household appliances + It is enough money for all except estate	4,3% (5,32)	4,3% (4,68)
There are no any problems with money	21,2% (26,09)	21,2% (22,91)
Total in group	100% (123)	100% (108)

Using the criterion χ2, it was shown that data of income higher than the "money is hardly enough for food" can be combined (χ^2 = 1,44, p = 0,98). Table 2x2 was the result. (Table 4). Analysis of these data showed (correct χ^2 20,41, p <0,0001), that CVD are significantly frequent in a group of women over 55 with incomes lower than "money is hardly enough for food" than in the group with more high incomes. At the same time there are significantly more women who have no cardiovascular diseases in the groups with higher incomes.

Table 4. Frequency of disease occurrence for women older
than 55 in groups with different income levels
(expected number are in brackets)

Income level	**There are** CVD	**There are** **not** CVD
"Money is hardly enough for food"	45 (29,82)	11 (26,18)
The level is higher then "Money is hardly enough for food"	78 (93,18)	97 (81,82)

The analysis of results has shown that for men, just as for women, frequency of Cardiovascular diseases significantly higher in group that evaluates their income as "money is hardly enough for food".

"Sick" people brings great economic damage. The data of this study showed that for the normal economic development is possible and

necessary to determine that minimum household income, which is applied least damage to health, that is, the residents have opportunities for health maintenance.

Conclusion

The increasing longevity of active life and corresponding increasing in expenses at public health services is a new call to world community which forces to search new ways of development of public health services and maintenance of active longevity. Researches, conducted in the developed countries, show necessity of using of information-communication technologies widely, to lower expenses and to raise quality of health services and to create conditions for comfortable life and work at elderly age. New market of services for elderly citizens has large potential consumption of advanced intelligent technologies.

At this stage, the problem of development and research is creation of system for gathering, store and editing data, obtained from quiestionaires of the elderly population of Republic of Kazakhstan. In the future functions of portal would be: questionnaire input by patients or their relatives; advanced statistical information from portal database questionnaires; ability to analytics; placement of scientific publications; scientific journal of the medical community; information exchange between patients, the portal and clinics, forum for portal users on gerontology or other medical issues. The preliminary analysis shows possibility of getting serious analytical results at stage of epidemiological screening. The expansion of research base with inclusion of block of clinical researches allows to expect that the portal becomes serious instrument of researches in the field of gerontology in Kazakhstan.

References
1. Международный (Венский) план действий по проблемам старения. резолюция ООН 37/51, 1982, http://www.un.org/russian/topics/socdev/ elderly/oldactio.htm
2. Принципы Организации Объединенных Наций в отношении пожилых людей. Сделать полнокровной жизнь лиц преклонного возраста. Приняты резолюцией 46/91 Генеральной Ассамблеи от 16 декабря 1991 года
3. Millennium Development Goals: progress towards the health-related Millennium Development Goals.Fact sheet N°290, May 2011. http://www.who. int/mediacentre/factsheets/fs290/en/index.html
4. Overview of the European strategy in ICT for Ageing Well. September 2009 (report). ICT for Inclusion, Directorate General for Information Society and Media, European Commission, BU31 01/66, B-1049 Brussels Belgium, http://ec.europa.eu/einclusion

5. National Geographic, January 2010

6. Ambient Assisted Living Roadmap. The Europian Ambient Assisted Living Innovation Alliance. Edited by Ger van den Broek (Philips Research Healthcare), Filippo Cavallo (Scuola Superiore Sant'Anna), Luca Odetti (Scuola Superiore Sant'Anna), Christian Wehrmann (VDI/VDE-IT). VDI/VDE-IT AALIANCE Office, Steinplatz 1, 10623 Berlin

7. ICT & Ageing – Final Report, January, 2010. edited by Lutz Kubitschke and Kevin Cullen (WRC), empirica Gesellschaft für Kommunikations- und Technologieforschung mbH, Oxfordstr. 2, 53111 Bonn, Germany

8. Long-term Care Challenges in an Ageing Society: The Role of ICT and Migrants. Results from a study on England, Germany, Italy and Spain. Authors: Stefano Kluzer, Christine Redecker and Clara Centeno, 2010, JRC Europian Comission, EUR 24382 EN

Authors:

R. I. Muhamedyev, PhD, professor, International IT University, IT department, Kazakhstan, Almaty, email: ravil.muhamedyev@gmail.com, phones: +7 701 3393720, +7 777 2399788

V. G. Dmitriyev, master of IS, International IT University, IT department, Kazakhstan, Almaty

A. T. Mansharipova, Doctor of Medicine, professor, Kazakh-Russian Medical University, Kazakhstan, Almaty

M. M. Maratov, master of IS, International IT University, IT department, Kazakhstan, Almaty

E. L. Muhamedyeva, master of SC, International IT University, IT department, Kazakhstan, Almaty

D. A. Shynybekov, professor, doctor of Economic Sciences, International IT University, Kazakhstan, Almaty

I. M. Ualiyeva, Candidate of Physico-mathematical Sciences, associate professor, International IT University, IT department, Kazakhstan, Almaty

D. V. Zagulova, Baltic Russian University, Latvia, Riga

MODERN PARADIGM AND TREND IN CLOUD COMPUTING

Danko Naydenov

Abstract: *Recently a new technology gaining popularity. This is the Cloud computing. The reason for gathering more and more supporters due to the advantages it offers and that are: full use of computer resources, increasing economic result through increased utilization, increase system reliability, lower power consumption.*

Key words: *Cloud Computing, IaaS, PaaS, SaaS*

To understand more easily what cloud computing represent there will be a parallel review between the on-premises and cloud application. For this purpose, first we will look at a typical structure of the on-premises application.

Each application can be considered to consist of three main layers:

basis – almost any application using an existing software for the machine on which it operates. These are usually operating system and its functions and standard libraries that are treated as local support. Historically this is the first stage in the development of computing. Examples of operating systems can provide Windows-based operating systems of Microsoft, as an alternative are different distributions of Linux, but for more complex computing machines it is used versions of Unix. Examples of systems for local support can be given .NET Framework technology of Microsoft or Java platform and with help of these two technologies can be developed from small and simple applications to large complex and WEB based applications.

infrastructure - with development of the computing some part of the application are carried on specialized machines. This process began in the 80s and 90s of twentieth century. As an example of infrastructure services can be given different types of databases. For those databases most often dedicated server on which they operate and other applications are accessing data through the network from another computer.

applications as services – with the creation of more complex applications most of them share some of their functionality that can be used by other applications. These are the current trends in computing. Applications that share its functionality can be considered as individually designed for very specific area and other developed by large companies with more wide scope.

Such described triple layer structure for the on-premises application was observed in cloud applications also. However, it should be noted that although there are similarities in structure of an on-premises and cloud application they are designed for different purposes, especially the way that they service users. In the on-premises application even it is Enterprise number of users is limited, while cloud application can be

expand greatly because of the fact that they are working in a Web environment.

As in the on-premises and in cloud applications basis provides basic functions for the operation of the application. A typical example of cloud operating system is Elastic Compute Cloud (EC2) of Amazon. EC2 is a modified Linux that runs on a virtual machine. Developers are free to install whatever software they want and libraries on this instance.

In infrastructure things are slightly different. If on-premises application choose to use Java, this does not limited it to use certain database, such as MySQL. In cloud applications this is not true. Most cloud providers offer own solution to store data and for that there is a good reason. This is scalability and to achieve this goal it is necessary providers to have the ability and freedom to optimize the system for storing data [4].

Cloud applications also share their functionality so it can be used by other applications. An example can be given to search engines that provide access to its search functionality so other applications can benefit from it. Another example of sharing functionality of cloud application is mapping. There are many activities, such as real estate, tourist attractions, shops, that want to show their location, but the developers either do not have enough budget or resources to make their own system of mapping therefore in such cases it is command to used cloud applications sharing such functionality.

The reason for which cloud computing is gaining popularity are several.
- Rate of load of computers [1]
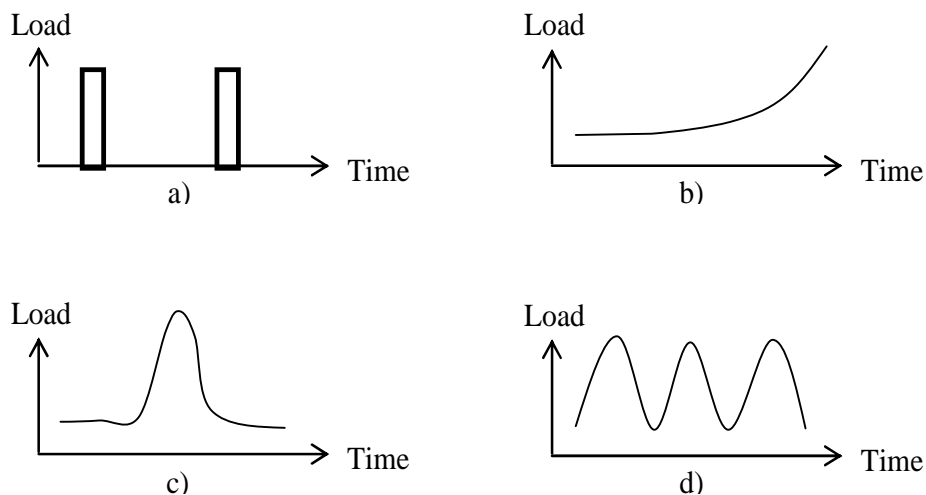
Figure 1: Computer load

233

In most cases, computers are loaded only at 10% -15% of the time. In the rest of the time they performs computing tasks, but only to supports the operating system. Example of this situation is a computers that runs batch job (Figure 1a).

Another situation where we have inefficient use of computing resources is when it is expected the load gradually increases over time and need to provide computing resources to meet this future workload. In this case it is needed a large initial investment and it is not known when exactly the resources will be fully loaded (Figure 1b).

A third situation in which there is a problem with the computing resources is when there is peak load at some point. An example is a large public event which will lead of big interest for information - such as elections for Parliament (Figure 1c).

The last situation which lead to inefficient use of resource is when there is recurrent load. An example of this situation is different sport events that take place every week (Figure 1d).

In all these situations there is inefficient utilization (idle) of computing resources. Cloud computing provide a good solution in this area. The key feature for them is that user pay only for used resources. In other words, the cloud environment applications can request CPU resource, use it and after completion the task the resource can be released and the payment is only for time and size of used resource. This feature of the cloud computing solves all of these problems because it does not require an initial investment of purchasing servers and there is not costs for their maintenance, while there is scope for expansion of the environment.

- In recent years there has been considerable development in computing, with its development increases the need for more space for storage, leading to increased costs. It also increases the energy consumption of computers and other peripherals.

- Striving to reduce the initial investment and subsequent maintenance costs of a technique such as server administration, maintenance of network devices, need for updating their software, problems as result from possible loss of data or problems with viruses. All these problems can also be solved with the usage of cloud computing.

Exact definition of CC is difficult to be given. This is an attempt for a brief description of that what is cloud computing.

Cloud computing is remote and distributed dynamic set of virtual computing resources available over the Internet, which provides an environment for deployment and execution of applications. In essence this is distributed computing environments from which applications can

234

request and use resources and after completion of the work the resource can be released.

Although it is difficult to give a precise definition of cloud computing there is agreement about what types of service are available [2]:

IaaS	PaaS	SaaS
User	User	User
Application	Application	Application
Data	Data	Data
Runtime	Runtime	Runtime
Middleware	Middleware	Middleware
OS	OS	OS
Virtualizati	Virtualizati	Virtualizati
Servers	Servers	Servers
Storage	Storage	Storage
Networking	Networking	Networking
a)	b)	c)
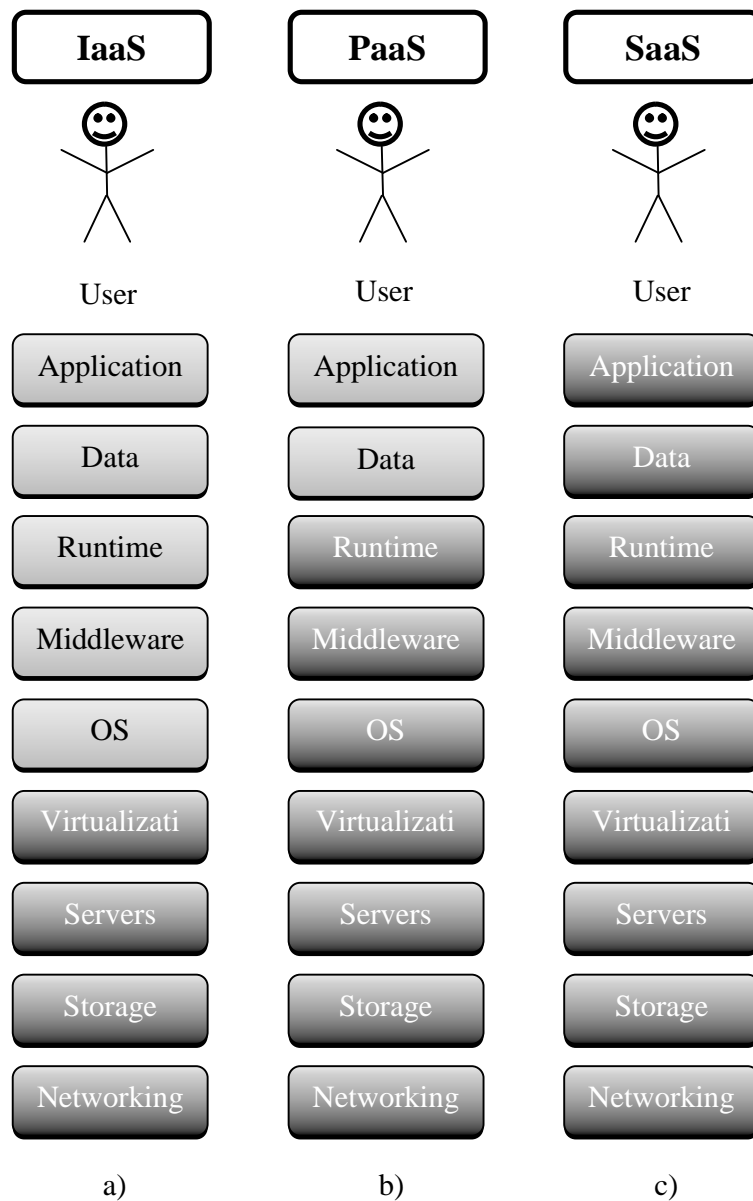
Figure 2: Layering of cloud service

infrastructure as a service (IaaS) – Customer of such type of providers normally gain access to different types of infrastructure. In general IaaS providers hold a very large physical infrastructure and this infrastructure is divided to smaller virtual resources which are accessible to customers. This type of service can be provides as raw virtual machine, or as a simple storage. Most completely solution is to be provide virtual machine(s) with operating system and possibility for multiplication. As it can be seen this type of service include layers from networking to virtualization (Figure 2a).

platform as a service (PaaS) – such providers commonly extend existing IaaS by appending middleware. The additional software representing this middleware commonly is Data base software, or some type of run time environments such as Java Runtime Environment (JRE) or a web server. Using this extension software more sophisticated applications can be started. This type of service include layers from networking to runtime (Figure 2b).

software as a service (SaaS) – Customers of such providers typically receive rich web interface. Normally users are completely separated from application that runs behind the interface. Providers of such type of service can easily inspect the load and if the additional calculation or data resource are required, the provider can increase the amount of machines in the cluster. This service include all layers from networking to application (Figure 2c).

There are several approaches that is used for developing and usage of cloud computing:

- In the first case cloud vendors provide an environment that is used by developers to create new applications. In this case direct users are developers, not end user.

- another option to use cloud computing is when an on-premises application uses functionality provided by the cloud. A typical example is iTunes of Apple. This is a desktop application used for playing music and movies, but also allows access to the cloud where it is possible to purchase movies and music.

- providing a complete solution of service from cloud provider is the last option. In this case end-user uses a small program on his computer or mobile device (typically a web browser) and accesses the service, which is located entirely in the cloud.

Unfortunately cloud computing has some deficiencies.

First is the issue of security. There are many cloud applications that store information about customer accounts, tax information, e-mails,

health data etc. Everyone is sensitive to such type of data and when these data are not under his control may appear different doubts[3].

Another feature which must be take into consideration is that cloud vendors charged usage of a resource. This means that whether the resource has been loaded or not when it was reserved user have to pay for it. Moreover if for some reason a resource is used inappropriate or unwanted (malicious access to the application) then it will again be paid.

Since cloud computing is a new technology and it is developing very dynamically all these deficiencies will find their solution in the near future .

Conclusion - In the field of computing it is rarely seen new technology. But when new technology is successful, it gives great impetus to the development of application. Cloud computing still can not fully replace the on-premises applications but they attract the attention of more people. The main reason is the ability to scale, high reliability and low cost for maintenance of the application. So if you are working in the field of software development you should expect that cloud computing will increasingly enter as a platform which will be used for creating an applications.

References
1. Dominique Bellenger, Jens Bertram, Andy Budina, "Scaling in Cloud Environments"
2. Frank Leymann, "Cloud Computing: The Next Revolution in IT" September 2009
3. Kaufman, L. M. "Data Security in the World of Cloud Computing" July 2009
4. Zhang Jian-hua. "Cloud Computing-based Data Storage and Disaster Recovery"

Author:
Danko Naydenov Naydenov, TU-Varna, 0899/066-331, danko@gbg.bg

DISPATCHING SYSTEM OF MINING AND TRANSPORT ENTERPRISE REQUIREMENTS SPECIFICATION

Ryskhan Satybaldiyeva, Aiman Moldagulova

Abstract: *DSMT (Dispatching System of Mining and Transport Enterprise) is used by the dispatcher station of a mining enterprise and is expected to use the hardware, operating system, software, and other resources available to that station. This article describes the customer's understanding of what resources are typically available in such a dispatcher station, and thus provides an understanding of the environment in which DSMT will operate.*

Keywords: *Dispatching system, On-board controller, Sensors weight, fuel, ignition, GPS-Antenna, Radio aerial, Radio station, The receiver of differential corrections, The Communication Computer, The server computer, dispatchers workstations.*

Introduction

Nowadays dispatching systems of mining and transport enterprises (DSMT) are successfully used in many pits of Kazakhstan. One of them is "Karier" system designed by VIST Group [3]. The common methodology of such systems based on imitation modeling. The integral part of the system is database which is based on the most powerful and reliable database platforms ORACLE. Dispatching systems are very important and sophisticated systems and allow to: 1) increase productive time for the use of equipment during working shift; 2) ensure resource saving and reaching the necessary volumes of production at the same time; 3) increase working and technical discipline of employees; 4) develop the basis for objective evaluation of different services and factory sectors activities; 5) create background for regular repairs and maintenance of a company's car park; 6) lay the foundation for the optimization of traffic [3].

The overall structure of the systems must provide continuous flow of information about work of transport equipment, conduct its analysis in order to ensure the efficiency of operation of machines. Obtaining the necessary information about work of transport equipment is made directly from on-board systems installed on transport or in the process continuous transmission of information or as a result of its reading from removable drives. This article deals with modern approach of software project development. We have paid particular attention to writing requirements specification as this stage defines 80% of the success.

General

Development of requirements should take into account the nature of the system. Essential considerations are emergent properties, the constraints and provisions of the external environment and the interfaces with surrounding systems [1]. DSMT is a complex multi-parameter dynamic system consisting of several subsystems (Figure 1) such as external environment, system of operational control and planning, system maintenance, transport equipment, locomotive train, mining dump truck, transport communications, railroad tracks, highways, cutter-loading equipment, shovel, place of loading, place of unloading [2].

The analysis of part of process develop is primarily concerned with understanding the nature and scope of the input requirements to assess the probable risks involved in satisfying them. Analysis work can range from feasibility studies to explore potential implementation options to the building of prototypes of some vital or high-risk components [1]. For understanding and determining the structure of derived requirements it used simulation. Usually for stakeholders requirements it is applied use cases or users scenarios, which help understand how user will exploit the system. In our system the interaction between user-dispatcher and the system is express by the model of user view given in the figure 2.
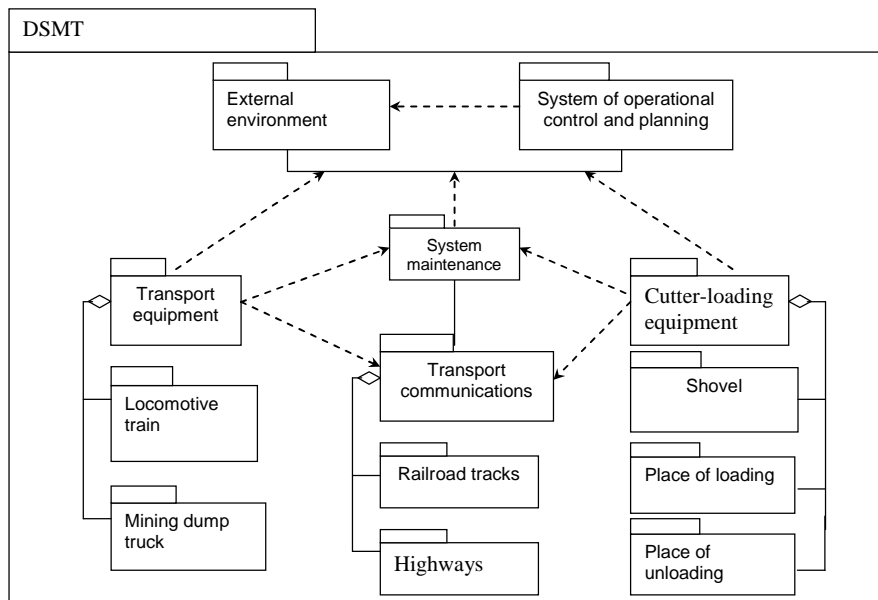
Figure 1. Diagram of packages (subsystems) of DMTS
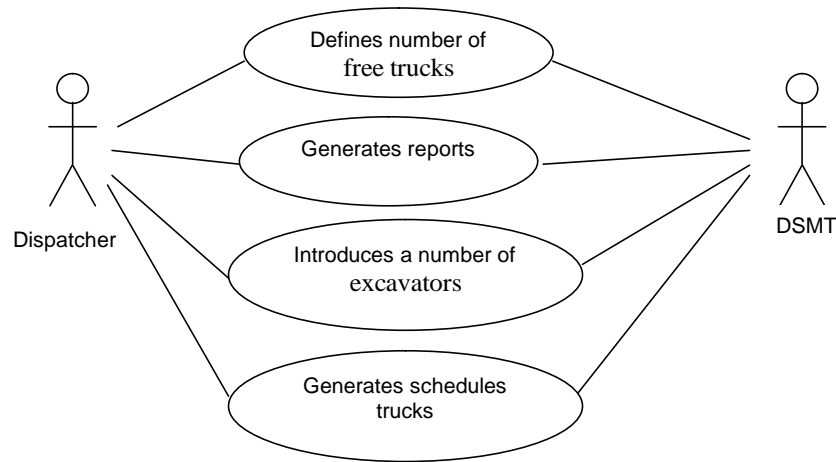
Figure 2. Use case diagram

During writing software requirements specifications it is selected an existing template that we could fine tune for our organizational needs [4]. Taking into account the recommendations given by Donn Le Vie and this consideration of existing standards we have created requirements in this article.

Basic questions that must be considered by an author of Software Requirement Specification (SRS), are as follows:

a) Functional facilities. What are the expected features of the software?

1. Monitoring delivery of ore from the pit to the point of admission

2. Calculate the optimal time for loading and unloading of cargo from trucks

3. Determine the optimal time interval between trucks

4. Provide a weekly report on logistics

b) External interfaces. How does the software interact with users, hardware systems, other hardware and other software?

1. Software shall be manage by the dispatcher

2. Check-out time and movement from a pit to the point of receiving sensors monitored

c) Performance. What are the performance, availability, response time, recovery time of various software functions, etc.?

1. The software operates without a break (or time)

2. Must be protected from failure. With the possibility of a power outage disaster recovery software and the latest data

d) Attributes. What are the portability, correctness, maintainability, security, software, and other criteria?

e) Design constraints imposed on the implementation of the product. Are there any required standards for the effective implementation of the language, the policy of preserving the integrity of databases, resource limits, operating environment (s), etc.?

A developer of SRS should avoid locating in the SRS requirements to the development or project [5].

Conclusion

As a result of the introduction of DSMT we will get an opportunity to solve problems of operational management of pits more effectively, including the problem of optimizing traffic, maintaining the required content of necessary amount of ore in warehouses, and also managing fuel filling. Moreover we could get a possibility for an objective assessment of services and enterprises, which has a positive effect on labor and technological discipline of staff.

References
1. Elizabeth Hull, Ken Jackson and Jeremy Dick. Requirements Engineering, Springer, pp. 201, 2005
2. Ахмедов, Д. Ш. Разработка научно-методических основ объектно-ориентированного моделирования геотехнических систем на карьерах – Диссертация, на соискание ученой степени доктора технических наук-Алматы. 2003.
3. www.vistgroup.ru
4. Donn Le Vie, Jr., Writing. Software Requirements Specifications (SRS), http://techwhirl.com/skills/tech-docs/writing-software-requirements-specs, 2010
5. http://www.webisgroup.ru/services/programming/srs/ieee-830-1998/

Authors:
Ryskhan Zh. Satybaldiyeva, Aiman N. Moldagulova
International Information Technologies University
Department of Information Systems and Mathematical Modelling
Manas Str./Zhandosov Str., 34 «A»/8 «A», Almaty, Kazakhstan,
+7 (727) 330-85-60; +7 (727) 244-51-21
E-MAIL: R.SATYBALDIYEVA@IITU.KZ, A.MOLDAGULOVA@IITU.KZ

REVISITING ABSTRACT SYNTAX TREE AS A BASIS OF SOURCE CODE KNOWLEDGE MODELS

Alexander Yanakiev, Haralambi Haralambiev, Kraicho Kraichev

Abstract: *Software modernization is saving costs in the development and maintenance of constantly evolving software systems. An abstract and suitable representation of the source code is a basis for performing the essential modernization techniques as well as source code knowledge models. This paper revisits AST and proposes practical guidelines for constructing a programming language independent entity-relational model for a representation of source code based on an AST structure. The software tool Smart Source Analyzer is presented as well, which implements these guidelines. This tool can be used as a basis for developing more complex and "smart" source code manipulation applications.*

Keywords:*static analyis, AST, entity-relational, source code modeling, tool implementation guideline*

Introduction

The significant problem in the industrial software development is that the developers do not always follow the best practices and standards for coding,adopted by software development team and this way produced source code does not reflect the project models and design decisions. There is an immense quantity of software systems and source code with tremendous commercial value.A considerable portion of it is challenging to modify and comprehend. Themain causes for this are projects' size, shortage of knowledge, old/new technologies, high cost of production, a lot of legacy code.

Application Modernization is a cost-effective way to preserve legacy investments in software and avoid the costs of migration. Software transformation is a way to achieve application modernization and as well is linked with solutions of problems caused by lack of knowledge of a software system. Suitable way of viewing the software system knowledge is building a visual representation of the system. In practice these modernization techniques are impossible to perform without a detailed model for representing and transforming the source code of the software system.

Related Work

Dynamic analysis is the process of evaluating a system on component based on its behavior during execution. *Static analysis* is the process of evaluating a system or component based on its form, structure, content, or documentation.It is a set of techniques for reading

the source code and analyzing it without execution **Error! Reference source not found.**. One of its purposes is finding defects and patterns andmaking estimations about what the behavior of the code will be at run time. This saves a great amount of resources as configuring an environment suitable for building and executing the code may prove difficult and expensive. For example, when testing every path in complex programs, it may be cheaper in terms of memory and processing power to perform static analysis, as dynamic analysis will require constructing a large number of test cases.

Models

Common software analyzing and modeling techniques involve parsing technologies and building a structure corresponding to the source code. Abstract Syntax Trees (ASTs) are formal representation of the syntactic structure of the code. They are used in compiler construction and can be viewed as intermediate data format of source code.

Theoretically the best situation will be if this intermediate data format is independent of the programming language and the target machine**Error! Reference source not found.** but in practice this is not the case. There are different Abstract Syntax Trees for different programming languages and even different Abstract Syntax Trees for one language. In spite of this they have several common properties. ASTs are language oriented – the structure of the tree node types is defined by language constructs and their semantic. They also contain information about the text of the source code – this is their main purpose though some are extended providing additional information.

An ASTmay be constructed using formal Context Free Grammar **Error! Reference source not found.** of the programming language. There may be different grammars for the same source language that produce similar but yet different ASTs.

Analysis may be performed over ASTs via traversing them in different ways. A suitable and common practice is the usage of the Visitor Pattern **Error! Reference source not found.**.

Elements of a software system may be viewed as elements of an oriented multigraph. Units of the software may be the vertices and dependencies between them may be edges of such a multigraph.

A detailed entity-relational (ER) **Error! Reference source not found.** representation of a software system is a suitable way to perform a thorough analysis on a software system. Representing the source code in that manner has the advantage that dependencies may easily be resolved even transitively as the model is simply a graph.

One common property of entity-relational models is obviously their structure which makes them abstract and complex. Entity-relational models are user defined and follow a specific goal. They are not part of any standard intermediate or object code building process. ER models may or may not be language independent.

Implementations

In recent years Java **Error! Reference source not found.** has become very popular for industrial, educational and research purposes.

There are different ways of obtaining an AST for code in the Java. One way to do it is to use a Lexer and Parser generating tool like ANTLR**Error! Reference source not found.**, JavaCC**Error! Reference source not found.**, YaCC**Error! Reference source not found.**, etc. A major challenge in this approach is to provide a complete and valid grammar to these tools for the programming language you are willing to parse. Programming languages fall into the class of context-free grammars, which makesbuilding such a grammar both a technically and theoretically demanding task.There is hardly a guarantee for its correctness and thoroughness. Moreover, the task is further complicated by the requirement that the grammar is suitable for constructing an AST. Figure 1 shows examples of two rules for defining a modifier and a variable initialize in the Java language.

```
modifier
    :  annotation
    |  'public'
    |  'protected'
    |  'private'
    |  'static'
    |  'abstract'
    |  'final'
    |  'native'
    |  'synchronized'
    |  'transient'
    |  'volatile'
    |  'strictfp'
    ;
variableInitializer
    :  arrayInitializer
    |  expression
    ;
```

Figure 1. Examples of rules in an ANTLR grammar

Another and a lot more secure way for getting an AST is using a frontend of a compiler tool that provides an API for using the AST, generated during compilation. Such APIs are the Java Compiler Tree API **Error! Reference source not found.**, provided by SUN Microsystems**Error! Reference source not found.**, and the Eclipse **Error! Reference source not found.** AST APIas a part of the Eclipse Java Development Tools (JDT)**Error! Reference source not found.**.

An important feature of frontends is binding trees from one to another, meaning from usage to declaration. Originally, an AST contains only syntactical information about matches of different rules of a grammar given source code as an input.

A compiler performs further analysis over the tree structures to collect more information about the source code. Important for this analysis are return types, tests whether a statement is reachable, etc. For example, a compiler must know what the return type of a method invocation is in order to test if the return value is assignable to the variable standing as a left expression of an assignment operator.

The Eclipse API offers the possibility to perform that additional analysis and resolve bindings between tree nodes. In the Compiler Tree API this is not straightforward as the user should himself perform analysis over the different data structures of the API providing richer information than the standard AST.

Entity-Relational Models

In **Error! Reference source not found.** a language-independent, scalable and extensible meta-model is defined. It is used in the development of X-Develop**Error! Reference source not found.**. Another similar work is mentioned in **Error! Reference source not found.**, wherean entity-relational Source Code Structural Model (SCSM) is defined for profiling of object-oriented source code. A major advantage of this model is its scalability, language variety, change propagation and others. The source code analyzers include parsers built with the JavaCC tool.

The Object Management Group (OMG) **Error! Reference source not found.** Application Driven Modernization (ADM) **Error! Reference source not found.** Task Force is developing a set of standards for modernization**Error! Reference source not found.**. These standards include KDM **Error! Reference source not found.**, ASTM **Error! Reference source not found.** and others:

KDM and ASTM are two complementary modeling specifications. KDM defines abstract semantic graph models and ASTM establishes specifications for ASTs. The primary purpose of KDM is to provide a

common interchange format allowing interoperability between existing software assurance tools, services, and their intermediate representations. From a meta-model perspective KDM is an entity-relationship representation.

All the models referenced above have several properties in common. They are all entity-relational. They all aim to be suitable for language-independent source code representation. KDM is a popular, open and established specification and an implemented model of that kind would be valuable for analysis.

As it is stated in the KDM specification:

"An entity is a thing of significance, about which information needs to be known or held. A KDM entity is an abstraction of some element of an existing software system that has a distinct, separate existence, a self-contained piece of data that can be referenced as a unit. Each KDM package defines several entity types representing specific abstractions related to a certain viewpoint on existing software systems.

A relationship represents an association, linkage, or connection between entities that describes their interaction, the dependence of one upon the other, or their mutual interdependence. A KDM relationship represents some semantic association between elements of an existing software system. Each KDM package defines several relationship types representing specific abstractions related to a certain viewpoint on existing software systems. All KDM relationships are binary."

Model Comparison

Choosing the most appropriate model is essential for performing analysis. When implementing a static analysis tool pros and cons of different modeling approaches have to be taken into consideration. An AST based model would be strongly language-specific. Such a structure is well defined and at times prone to modification but rarely to extension and enrichment with additional knowledge. An entity-relational model may be designed as language independent but thus becoming highly abstract and complex. Parser and compiler tools do not normally provide such a representation like they do with the AST representation and additional analysis should be performed.

The degree of extensibility of a user-defined entity-relational model depends on the model's design. The model, defined in the Knowledge Discovery Metamodel specification is highly extensible and designed to be language independent.

Proposed Approach

Ability to analyze software system artifacts in a language independent manner is critical for the analysis of the modern software systems, which are usually complex and heterogeneous – using different technologies and programming languages. Taking into account that AST modeling is strongly language oriented we decided instead to create and use a KDM based entity-relational model using the AST. KDM model is language independent in its nature as discussed above.

Even though there is no standardized AST interface to use for the implementation of an entity-relational model similar to this one KDM specifies, there are common guidelines that apply to the most of the front-ends of the compiler technologies. These generic rules derive from the semantics of different syntactic structures in the programming language and may be applied to any AST.

Tree Node Commons

The first common feature in various implementations of syntactic segmentations is the division of statements and expressions. The basic semantic difference between them is that expressions evaluate and statements do not. Statements may contain expressions but the opposite is not valid.

A statement might be a block, a loop, an if-else, etc, whereas an expression might be a method invocation (which itself might be contained in another expression or an Expression Statement), an identifier, binary, unary, etc.

Different statements or expressions have different contents – blocks contain a list of statements, switch statements contain a list of cases (all of them have an expression except for one optional – the default), a binary expression contains two expressions - one for each side. A specific handler object should be applied to different tree nodes (expressions, statements, etc.) but implementing this specific handler for a different AST should be similar as are the semantics of the tree nodes.

Tree structures are provided in a way that a visitor pattern may be applied in order to obtain further information about nodes contained in the current node or the nodes dependencies.
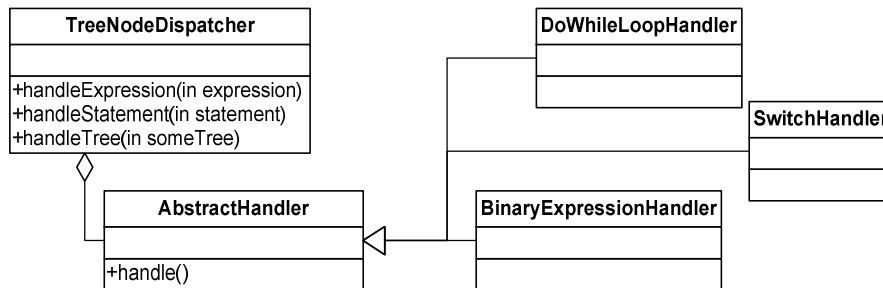
Figure 2. A UML diagram of the Dispatcher – Handler relation

Algorithms for implementation

KDM is structured into layers which are segmented into packages. Packages define sets of meta-model elements that stand for a certain independent knowledge facet of the software system. There is a model in the package referring to all the entities. An exception to the rule is the action package. There is no action model as all action entities are code elements and may be obtained via the code model.

Several passes are needed in order to initialize the inventory and code model of the KDM representation. The inventory model is created with analyzing source and binary files contained in the project, the project's classpath.

The most important aspects of the representation areownership, model relations, and relations with other models.

Ownership is a powerful way of segmenting entities and introducing a hierarchy among them and their relations. If one entity is being owned by another, then the first one derives from the other. A component performing analysis upon a model of that kind may or may not choose to examine the contents of an entity in deep depending on the analysis' goals. Ownership is propagated to all views of the software system, i.e. all models. Traversing owned elements with an entity considered as a root results in an ownership tree. Examples of ownership in the Inventory Model are directories containing source files or binary files (the directories are the owners).

Creating the inventory model consists of traversing files and resolving classpath dependencies between the project and binaries or other projects in order to implement inventory model relations.

Code model implementation consists of multiple traversals of the AST and related data structures provided by the compiler front end used for model creation. Examples of key relations to the model we use are class hierarchy ones like *extends*, *implements*, *overrides*, etc.

248

One of the features our implementation adds to the original KDM specification is the return types and return data elements of actions. The difference between statements and expressions is that expressions may be evaluated and may have a return value or at least type. An *instanceof* operator in Java has an expression in its right side which evaluates to a type.

The return object may be a type when the expression is of type identifier in a *type cast* and the identifier resolves to a Type.

Some actions, representing statements, are also assigned a return data element.For example, a *Variable Declaration* is assigned the created variable as a returned data element.

Return data elements are useful to calculate a tentative number of created temporary data elements during execution. They are also useful for obtaining the return object of an action that may be written into another variable for further data flow analysis. These data elements however may be distinguished and thus excluded from the user by its kind. These temporary data elements are marked differently from all others.

Action model building examples

Our implementation of the entity-relational representation consists of several directed multigraphs for several different views (called *models*) of the software system. There are numbers of various types of relations that may exist between vertices from the multigraph standing for software elements with different semantics. There are also relations from one multigraph to another thus resulting in one multigraph providing means to access information about all important aspects of the software system directly.

Building the action model for a block is an easy task taking into consideration the semantics of a block in a programming language. First of all, the block of the AST Node for the method is obtained via the compiler front end. Semantically a block is nothing more than a list of statements. There may be branching inside them but not between them.

Figure 2 shows a basic and high level design of a system dedicated to transforming an Abstract Syntax Tree, or at least using it as a read-only source, into an entity-relational model. There are two main types of roles in the process – a dispatcher and handlers.

A dispatcher is an object that acts as a standard interface for any other object that wants to interpret an AST node. The dispatcher receives a node and
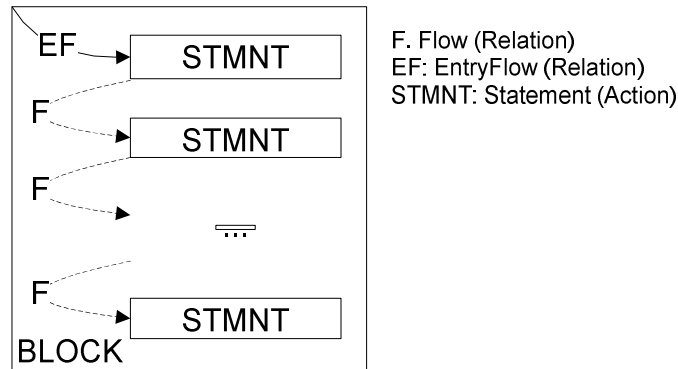
Figure 3. An example of how entities and relations are linked
in a block action

After the parsing of a method has started, its block is passed to the Dispatcher, which uses one concrete *Handler* (a *block handler*). According to the block element's semantics the statements have to be analyzed and flow relations have to be created between them. This is achieved by passing each statement of the collection to the *Dispatcher* which then uses another *Handler*(see Figure 2) to create an *ActionElement* representing the statement. The *Handler* objects are responsible for creating fully functional action elements as is the *block handler* itself (it creates a fully functional element in the model – a block unit). After the action elements are created they are returned to the *block handler*. The *Handler* has to create the relations between the entities owned by the element it creates. It does not create the relations between the element it creates and its siblings in the ownership. There are *Flow*-relations created between the consecutive actions in the list. There is an *EntryFlow* relation between the block and the first owned action. Figure 3 shows an example of what a *block action* looks like.

Another example of a handler is a *switch handler*. As stated above a switch contains a collection of cases.

Figure 4 shows an AST constructed for the following Java code:

```
switch (dataElement)
{
case 2:
  int x = 1;
case 4:
  x = 2;
  System.out.println(x);
}
```

Figure 4. An example of an AST, corresponding
to a switch-case construction

When a Switch enters the *Dispatcher* it is processed to a *switch handler*. The handler gets all the cases and sends them back to the *Dispatcher* which forwards them to a *case handler*. The case has similar semantics to a block as it contains a collection of statements but the difference is that it does not close a scope. This means that a variable declared in a case is valid until the end of block the case is defined in. Another difference is the case's expression which may be an action or a data element and there is a Reads relation from the case action to that data element.

Firstly, the *switch handler* takes care of the switch expression. The expression of the switch may be an action if it is for example a method invocation or just a data element if there is a variable. The expression is again sent to the *Dispatcher*, processed by it and then returned as a *DataElement* or *Action*. If it is a *data element*, it is wrapped up in a special Action with a micro kind property *data action*, which is another addition to the original KDM specification. The point of that is to have an *entry flow* relation between the whole action representing the switch and the action for the switch expression.

F: Flow (Relation)
EF: Entry Flow (Relation)
GF: Guarded Flow (Relation)
FF: False Flow (Relation)
C: Case (Action)
D: Default (Action)
SE Act: Switch Expression (Action)
SE RDE: Switch Expression Result (Data Element)
R: Reads (Relation)
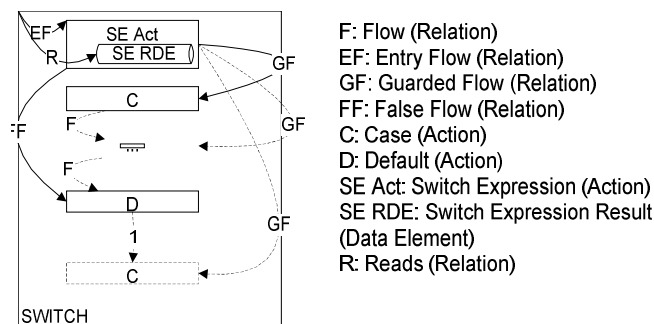
Figure 5. An example of how entities and relations are linked
in a switch action

After the *case handler*objects have created action elements for the statements in the case-sthey return the created action elements to the Dispatcher which forwards them to the *switch handler.*The *switch handler* receives all the actions representing the case-s and creates flows between every two consecutive cases where the first one does not contain a break. Furthermore it creates *guarded flow* relations from the action for the switch expression to the case actions and a *false flow* relation to the default action (if such action exists). Figure 5 shows how a switch statement might be presented in the entity-relational format.

Examples

Here are several examples of what different actions might look like provided with the sample source code. The examples show how expressions might be processed as the types of the expressions may vary.

The current example shows three different ways of constructing a return statement and its expression. Figures6and 7 show how a return statement's expression may be an expression that is an action. The concrete example of this action is a method invocation but it may be a ternary operator, or even a field selection with the '.' operator.

```
public class M
{
    private intrInt()
    {
        return 3;
    }
    public intreturnFunction()
    {
        return rInt();
    }
}
```

Figure 6. An example of a return statement, returning the result
of an action element

R: Reads (Relation)
RE Act: Return Expression (Action)
RE DE: Return Expression Data Element *Data Element)
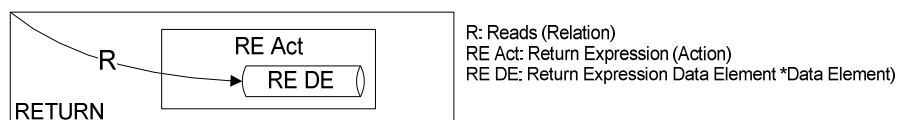
Figure 7. A graphical example of the entity-relational representation
of the example shown in Figure 5
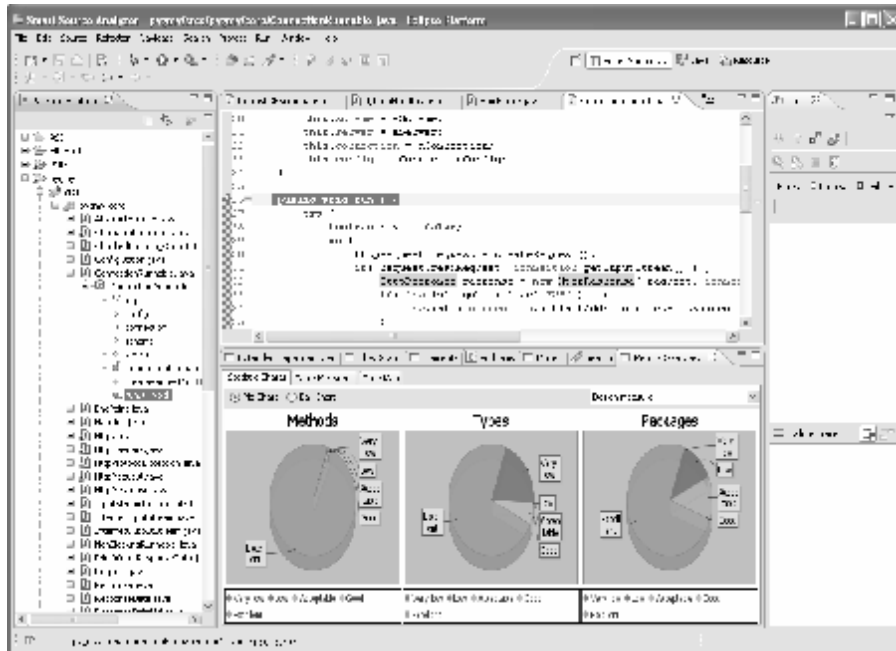
Figure 8. A snapshot of the Smart Source Analyzer integrated
in the Eclipse IDE

Prototype

The approach defined above is implemented in the Smart Source Analyzer (SSA) project **Error! Reference source not found.**. The SSA is an intelligent platform for modernization and development.It measures different aspects of a software system, looks for patterns and anti-patterns, performs automated generation of test sets. All these modules depend on one common model implemented according to the algorithm above.Figure 8 shows a snapshot of the SSA product.

The guidelines proposed in the paper serve as an algorithm for implementing one of the main SSA components. This component is responsible for reading the source code, analyzing it and disposing it to other components for performing further and more specific analysis.

There are two examples presented below showing calculating the Depth of Inheritance Tree (DIT) metric **Error! Reference source not found.** and looking for the Circular Dependency anti-pattern **Error! Reference source not found.**.

Depth of Inheritance Tree (DIT)

The Depth of Inheritance Tree metric is defined as the maximal length from the class to the root in the inheritance graph. A variation of

this metric is to remove the interfaces from the tree since the class cannot inherit functionality from then and thus such inheritance does not increase its complexity.

As the representation of the software system is an entity-relational model there are entities in that model corresponding to classes in the Object Oriented source code. The classes are entities and inheritance is a relation between them. *Extends* and *Implements* relations are part of the relations that might be owned by an entity According to the KDM specification. Obtaining only these relations of an entity (a ClassUnit or InterfaceUnit) and continuing the same process for the results will create a subgraph in the entity-relational model. Traversing this graph and calculating the number of edges to the most distant vertex results in the value of the DIT metric. Figure 9, 10 and 11 show how an inheritance tree is represented in an entity-relational model and a UML diagram of it.

UML representation
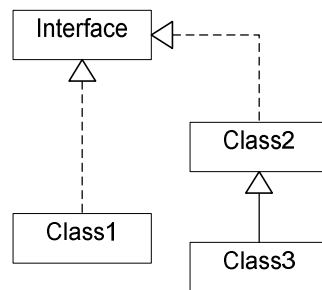
Figure 9. A UML Diagram showing an interface and three classes

Entity ID	Entity Type
Interface	InterfaceUnit
Class1	ClassUnit
Class2	ClassUnit
Class3	ClassUnit

Figure 10. A table showing instances of entities created in the model
for the types shown in Figure 9

Relation ID	From	To	Type
Impl1	Class1	Interface	Implements
Impl2	Class2	Interface	Implements
Impl3	Class3	Class2	Extends

Figure 11. A table showing instances of relations created in the model between
the entities in Figure 10 representing type dependencies from Figure 9

254

Circular Dependency anti-pattern

Circular dependencies might cause many undesired effects in software. The most problematic of them is tight coupling, meaning components are not only linked together but also depend on each other. This might cause a domino effect when a change should be introduced in one of the coupled components.

The model described in the paper is a directed multigraph. A pattern recognition component may use it to find different dependencies between vertices of that graph. Such a dependency is the Circular Dependency anti-pattern.

An exemplary approach to finding that pattern is specifying a subset of all different relations, defined in the model. A predicate for checking whether there is a circular dependency between two software components should be equivalent to one checking whether there is a path in the directed multigrahs from the first entity to the second and vice versa. In that way the client component of the model may specify different subsets of relations for different checks for that anti-pattern.

Conclusion and Future Works

An extensible entity-relational model is a suitable way for representing source code and its features. Building a language independent model of that kind would provide a standard interface for analysis components to evaluate software systems written in different languages.

In this paper, we presented a guideline for transforming a standard source code model into an entity-relational model with the purpose of being language independent. The algorithm shows a general approach that may be applied for various standard frontends for different languages providing a similar in structure Abstract Syntax Tree.

The main goal of building a language-independent software model is performing language-independent analysis over it. There is another interesting challenge originating from this approach – automated translation. This paper shows an algorithm for building a model from source code. The opposite algorithm is also required for automated translation – generating the source code from the abstractions – the entities in the model.

Another challenge which is connected to automated translation is serialization. Implementing the algorithm stated above for different languages requires the use of different parser/compiler tools. These tools are mostly frameworks in the target language analyzed by them so communicating between objects created by one framework for one language and another framework for another language might prove

difficult. Storing the model in an intermediate format would avoid these problems and would gain the possibility of other technologies accessing the models.

As software systems grow, building such rich models may require a large amount of resources and time. Optimizing build techniques should be an important feature of a prototype with these capabilities. One way of achieving this is incremental model building and change propagation. There are compilers which work only on changed resources but as the model aims to be language independent it should not keep its links with the technology used for building it.

Combining these features would provide the advantages of this modeling in all stages of software development.

References
1. Chen, P. (2002) Entity-relationship modeling: historical events, future trends, and lessons learned. Software pioneers: contributions to software engineering book contents
2. Chidamber, S., Kemerer, C. (1994) A metrics suite for object oriented design. IEEE Transactions on software engineering, vol. 20, no. 6
3. Chomsky, N. (1956) Three models for the description of language. I.R.E. Transactions on Information Theory IT-2.
4. Dereulle, L., Melab, N., Bouneffa, M. (2007) Analysis and manipulation of distributed multi-language software code. Proceedings of the 11th IASTED International Conference on Software Engineering and Applications
5. Fischer, G., Lusiardi, J., Wolff von Gudenberg, J. (2007) Abstract Syntax Trees – and their Role in Model Driven Software Development. Proceedings of the International Conference on Software Engineering Advances
6. Gamma, E., Helm, R.,Johnson, R. (1995) Design patterns elements of reusable object oriented software. Addison Wesley
7. Strein, D., Kratz, H., Löwe, W. (2006) Cross-language program analysis and refactoring. Proceedings of the Sixth IEEE International Workshop on Source Code Analysis and Manipulation
8. Strein, D., Lincke, R., Lundberg, J. (2006). "An Extensible Meta-Model for Program Analysis", Proceedings of the 22nd IEEE International Conference on Software Maintenance table of contents
9. ADM Standards Roadmap, http://adm.omg.org/ADMTF%20Roadmap.pdf, February 2009
10. Another Tool for Language Recognition, ANTLR, http://antlr.org/

11. Circular Dependency anti-pattern, http://en.wikipedia.org/wiki/Circular_dependency
12. EclipseJava Development Tools, http://www.eclipse.org/jdt/
13. IEEE, "IEEE Standard Glossary of Software Engineering Terminology,"
14. Java Compiler API, http://jcp.org/en/jsr/detail?id=199
15. Object Management Group (OMG), http://www.omg.org/
16. OMG, Abstract Syntax Tree Metamodel (ASTM), http://www.omg.org/spec/ASTM/1.0/Beta1/PDF/
17. OMG, Architecture-Driven Modernization (ADM), http://adm.omg.org/
18. OMG, Knowledge Discovery Metamodel (KDM), http://www.omg.org/spec/KDM/ 1.0/, January 2008
19. Smart Source Analyzer (SSA) – a product for static analysis, developed by Musala Soft - http://www.musala.com/
20. Sun Microsystems, acquired by Oracle – http://oracle.com/sun
21. The Eclipse Integrated Development Environment – http://eclipse.org/
22. The Java Compiler Compiler, JavaCC, https://javacc.dev.java.net/
23. The Java Programming Language, http://java.sun.com/
24. Yet Another Compiler-Compiler, Yacc, http://dinosaur.compilertools.net/

Authors:
Alexander Yanakiev, Musala Soft, alexander.yanakiev@musala.com
Haralambi Haralambiev, Musala Soft, haralambi.haralambiev@musala.com
Kraicho Kraichev, Musala Soft, kraicho.kraichev@musala.com

A MODEL OF DISTRIBUTED TWO-PHASE LOCKING WITH TIME STAMPS IN DISTRIBUTED DATABASE MANAGEMENT SYSTEMS

Branimir Zhekov, Svetlana Vasileva

Abstract: *The paper views the problems, connected with simulation modeling of concurrency control of distributed transactions in distributed database systems. There is a two phase locking simulation model of distributed transactions with using the timestamps ordering method for preventing deadlocks. The model is developed with the tools of simulation modeling environment GPSS World Personal Version. The distributed databases modeling in partial replication in concurrent upgrade of data elements copies is conducted in defined and accepted restrictions. The results from the simulations are presented and a comparative analysis with models of two-version two-phase locking is made.*

Key words*: distributed transactions, distributed databases, concurrency control, simulation models, time stamps, GPSS blocks*

The paper is dedicated to the developed algorithms for concurrency control of transactions in distributed database management systems (DDBMS). The task for comparing the different methods for the parallelism management of the transactions in distributed database systems (DDB) has not been investigated enough so far. There is almost a lack of results from simulation modeling of the methods using locking (*two-phase locking - 2PL*) in DDB. In [5] we have presented the results from the comparative analysis of algorithms *Ordering Network* and distributed *2PL*, in a condition of complete replication of on the nodes of DDBMS. And in [6] we have considered and analyzed one version and multi version algorithm but it is about Timestamp ordering method, where the transactions rollback more often then the 2PL protocols.

In the contemporary DDBMS the concurrency control of the transaction is based on the 2PL protocols. The data in such information systems are not completely distributed along the system nodes. The development of an analytical model for such systems is practically impossible for the following reasons:

partial replication of data along the nodes of the distributed system,
presence of restarts of transactions
and in some protocols [5] – presence of multilevel data architecture.

The features that are pointed out, give reasons for investigation of the productivity of DDBMS that work according to the method of *2PL,* and also reasons the method of simulation modeling to me chosen. The development of simulating algorithms and models and the conducted simulations are executed in the environment of GPSS World.

Topical for the development and the modeling algorithms with mechanism for deadlocks (*deadlocks* – DL) preventing: version algorithms for transaction management with locking, the method *time-out* and timestamp ordering method.

The simplest method for deadlock avoiding, according to a number of authors, [4], [5] et al., is the method of *time-out*. But the disadvantages of this method are: there is no guarantee, that the restarted along *time-out* transaction had taken part in the deadlock; with the increase of the time limit the possibility that the transaction is involved in the deadlock also increases, the time for discovering deadlocked transactions increases, and the time for the stay of the transactions that wait for the release of resources from deadlocked transactions increase respectively.

The *timestamp ordering* method has two basic advantages: no deadlocks are admitted ([4], [5] and others) and it is simpler for realization. No cyclic restart of a cancelled transaction is allowed, because during the restart, the timestamp is saved and every transaction as time passes by, is going to become "the oldest one" and consequently it is not going to be restarted in a competition for the resource [4]. The common (general) case of timestamp ordering algorithm is the one that follows: if the transaction that took the element is "younger" than the applying one, the applicant is waiting for the release of the data element; if it is older, the applicant transaction is restarted.

We consider the algorithm that models the execution of global transactions by the two-phase locking protocol, in which time stamp ordering mechanism for deadlock preventing is included. Every data element has two replicas. The distributed *2PL* protocol is viewed in details in [1] and [5] and the principles of timestamp ordering mechanism for deadlock avoiding – in [4] and [5].

The execution of global transaction is demonstrated in Fig. 1. The transaction is initiated from the node S_{P2} and it updates the data elements *El1* (that have replicas in the local databases LDB_{P6} and LDB_{P7}) and *El2* (with replicas in the local databases LDB_{P8} and LDB_{P9}), on the nodes S_{P6} and S_{P7}, and S_{P8} and S_{P9} respectively. In this model there are generated 6 streams of GPSS transactions that imitate global transactions in DDBMS. Each stream has intensity • (number of transactions/millisecond). Such a model of distributed transactions in DDBMS by GPSS transactions is described in [3]. In this paper and in [3] we consider the execution of transactions 1 and 2 data elements long. The possibility for short or long transactions occurrence is defined by the function *BrEl*.

We view the sequence of operations in the scheme which is presented in Fig. 1.

When the transaction T_{P2}^{P1} comes into transaction manager TM_{P2} its length is checked (1 or 2 data elements will be processed) – operation 1 and the GPSS transaction is repaired to splitting – operations 2. With operations 3 the values of the sub-transactions parameters are converted – the numbers of the lock managers LM_{P6}, (LM_{P7}), (LM_{P8} and LM_{P9}), where the sub-transactions $T_{P2,P6}^{P1}$, ($T_{P2,P7}^{P1}$), ($T_{P2,P8}^{P1}$ and $T_{P2,P9}^{P1}$) have to receive the lock for read/write replicas of data elements $El1$ and $El2$. In the common case it is executed transferring of requests for locking data elements replicas through network to the executor nodes (operations 4 and 10).

In the executor nodes S_{P6}, (S_{P7}) the lock managers LM_{P6}, (LM_{P7}) check in the lock tables LT_{P6}, (LT_{P7}) with operation 5 in Fig. 1 the possibility for presenting the locking of the replicas of the element $El1$ to the sub-transactions $T_{P2,P6}^{P1}$, ($T_{P2,P7}^{P1}$). The decision for presenting a locking of an element is accepted by the lock managers LM in conformity with the table of the compatibility of the locking shown in [1] and [4]. If the locking is allowed (operations-messages 6) the sub-transactions are split and their heirs (with operations 7) come back in the node-initiator for transmitting the confirmation for the locking of $El1$ before the sub-transactions $T_{P2,P8}^{P1}$ and $T_{P2,P9}^{P1}$ so that it would be possible that the global transaction continues its first "expanding" phase, and "the parents" $T_{P2,P6}^{P1}$, ($T_{P2,P7}^{P1}$) continue to execute the operations read/write (operations 9) on the replicas of the element $El1$ in the local databases LBD_{P6}, (LBD_{P7}). In most cases the locking of the copies of the element $El1$ is submitted.

Receiving the confirmation for locking of the element $El1$ (operation 8 on fig. 1) the transaction manager T_{P2}^{P1} - TM_{P2} transmits the sub-transaction that processes the element $El2$ to the transaction coordinator TC_{P2}. There is a great probability that (from the range of 0,8) the transaction updates $El2$, and therefore the sub-transaction is split (operation 10) and the sub-transactions $T_{P2,P8}^{P1}$ and $T_{P2,P9}^{P1}$ are transmitted through the channels of the network the corresponding lock managers LM_{P8} and LM_{P9}. LM_{P8} and LM_{P9} through the lock tables check the possibility for taking the replicas of the element $El2$ (operations 11). Submitting the locking (operations 12) the sub-transactions continue to the data managers DM_{P8} and DM_{P9} for execution of operations read/write of the element $El2$ (operations 14 on fig. 1). If the locking is impossible it is checked if the number of the sub-transaction is smaller than the number of the sub-transaction that has put the locking: if the sub-transaction does not continue and is not going to then it queues before the replica of $El2$ in the lock tables LT_{P8} and LT_{P9}. The waiting is modeled by user chains with number $P11$ (of the parameter $P11$ of each of the sub-transactions $T_{P2,P6}^{P1}$, ($T_{P2,P7}^{P1}$), $T_{P2,P8}^{P1}$ and $T_{P2,P9}^{P1}$ is given a value

before it enters the corresponding lock manager); if the sub-transaction has not received the locking and is not going to wait for its submission, it is restarted (operations 13 / operations 15 in Fig. 1). After arriving in TM_{P2} (operation 16), the restarted transaction is transmitted again to the lock managers (operations 17 in Fig. 1). The second (successful) attempt for locking the element 1 / element 2 is shown with the operations 18 and 19. The corresponding lock managers LM put the record for the element lock in the lock tables. After the execution of operations read/write (operations 9 and 14), the locking of the replicas of the elements is released (operations 20 – request and operations 21 – confirmation for removing the locking).
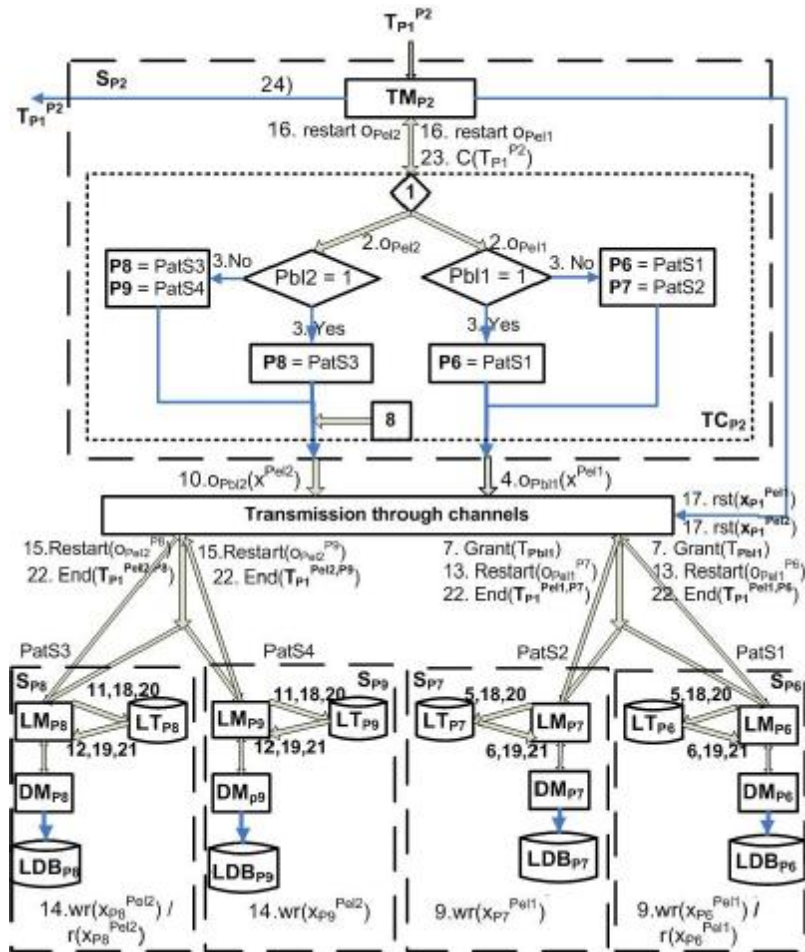
Fig. 1. A frame scheme of global transaction execution by the algorithm of Distributed 2PL with timestamps (in two data elements update)

The confirmation about the finish of reading/writing of the element $El2$ is transmitted to the transaction manager TM_{P2} (operation 22). The sub-transactions $T_{P2,P8}^{P1}$ and $T_{P2,P9}^{P1}$, and before that $T_{P2,P6}^{P1}$ and $T_{P2,P7}^{P1}$ (if the element $El1$ has been updated) are merged in the sub-transactions processing $El2$ and $El1$ respectively. The confirmation about the end of the corresponding sub-transactions is transmitted to the transaction manager TM_{P2}. The sub-transactions that processed $El1$ and the element $El2$ respectively, and the parent-transaction, waiting for them in the transaction manager TM_{P2} are merged (operation 23). After gathering the necessary statistics about these which have finished their work GPSS transactions, they leave the system (operation 24).

The following parameters for modeling transactions of DDBMS transactions are used in the modeling:

$P1$ – number of the guarantee transactions;

$P2$ – number of the generating transaction site;

$Pel1$ – number of the first processed data element by the transaction ($El1$);

$Pbl1$ – type of the requested locking of the element $El1$: 1 (rl) – if read ($El1$); 2 (wl) – if write ($El1$);

$Pel2$ – number of the second transaction processed data element ($El2$);

$Pbl2$ – type of the requested locking of the element $El2$: 1 (rl) – if read ($El2$); 2 (wl) – if write ($El2$);

$P5$ – phase of the transactions processing: it takes the value of 0 in the transaction coming in the model and after the end of the operation read/write it takes the value of 1.

$P6$ – number of the site where the first copy of the first data element $El1$ is stored;

$P7$ – number of the site where the second copy of the first data element $El1$ is stored;

$P8$ – number of the site where the first copy of the second data element $El2$ is stored;

$P9$ – number of the site where the second copy of the second data element $El2$ is stored;

$P11$ – number of the user's list where the corresponding sub-transaction waits for the release of the copy data element.

Simulation

Some fragments of a program GPSS World code, specific for the distributed transactions modeling and their service in embedding of the timestamps mechanism in the Distributed 2PL, are given on fig. 3. We

view a model of Distributed 2PL with timestamps synthesized similar to the modeling algorithms in [3].

```
…
LockTip1 FUNCTION RN3,D3        ;Probability of coming up of the
.40,1/.70,2/1.0,3              ;particular type of El1 locking
LockTip2 FUNCTION RN3,D3        ;Probability of coming up of
.20,1/.50,2/1.0,3              ;the particular type of El2 locking
DistrS1 FUNCTION V$SiteRepl1,D6     ;Replication of the data
1,2/2,6/3,1/4,5/5,3/6,4            ;First node for the El1
… ;Replication of the data – Second node for the El1
...
BrEl FUNCTION RN4,D2            ;Number of the elements accessed transaction
.30,1/1.,2
; Description of the used matrices
; Description of the storages
; Description of the tables for statistic
*** Segment for transaction generation
        GENERATE 60,FN$XPDIS    ;Generation of GPSS transactions with the given incoming
Potok1  ASSIGN 2,1               ;interval (60 ms)
        ASSIGN 1,(MP2+FN$NomSait)
        TRANSFER  ,BEGI
…
        GENERATE 60,FN$XPDIS   ;Generation of GPSS transactions
Potok6  ASSIGN 2,6
        ASSIGN 1,(MP2+FN$NomSait)
BEGI  GATE FV FN$TraMan         ;Processing in transaction manager
        SEIZE FN$TraMan
        FUNAVAIL FN$TraMan
        ASSIGN Nel,FN$BrEl              ; Giving the number of the elements
        ASSIGN El1,V$ElemN1            ; Giving the number of the El1
        ASSIGN Bl1,FN$LockTip1           ; Lock type of El1
        ASSIGN 5,0                     ;Flag of transaction committing
...
        SAVEVALUE BROITR+,1       ;Count generated transaction
        TABULATE TablGen
        TEST E P$Nel,1,Prebro2
        SAVEVALUE BROITR1+,1           ;accessed only one element
        TRANSFER ,Prebrg
Prebro2    SAVEVALUE BROITR2+,1   ;accessed two elements
Prebrg     ADVANCE 4
Obrab     QUEUE FN$TransCor
        GATE FV FN$TransCor             ; Processing in Transaction coordinators
        SEIZE FN$TransCor
; Preparation for (not)splitting of transaction to sub-transastions
; Defining values for P6, and for P7,
; and for P8, and P9
Pered2   RELEASE FN$TransCor
        FAVAIL FN$TransCor
        SPLIT 1,Pat1,REPLI             ;Splitting of the first subT
        TEST NE PBl1,1,Per2
        SPLIT 1,Pat2,REPLI
Per2   TEST NE P$Nel,1,PrKrai
Pat1ob   TEST E PBl1,1,Per22
        GATHER 2
        ASSEMBLE 2
        TRANSFER ,Per24
Per22   GATHER 3               ; The second sub-T waits lock(El1) confirmation
```

```
        ASSEMBLE 3
Per24   SPLIT 1,Pat3,REPLI      ; Spliting of the second subT
        TEST NE PBl2,1,PrKrai
        SPLIT 1,Pat4,REPLI
PrKrai  TRANSFER ,Krai
******************************************************************************
Pat1  ADVANCE MX$RAZST(P2,P6),MX$RAZDEV(P2,P6) ; Transfer in the net to
LockR1   ASSIGN NoEl,1                                  ; the LM of first copy of El1.
        QUEUE P6                        ; Waiting in front of LM of first copy of El1
        SEIZE P6                        ; Processing in Lock manager of first copy of El1
…
        TEST E P5,0,Fiksira1
        TEST E CH*El1,0,Chakane1              ; Is El1 first copy free,
Zaema1  TEST E MX$LTA1(P$El1,1),0,ProvR1  ; If "YES" then
        MSAVEVALUE LTA1,P$El1,1,P$Bl1    ; the element El1 first copy
        MSAVEVALUE LTA1,P$El1,2,P1       ; from lock table is taken
        MSAVEVALUE LTA1,P$El1,3,P2
        TRANSFER ,PatPot1
ProvR1  TEST E V$RAZRBL1,1,Chakane1  ; if El1 first copy is busy by
        ASSIGN CHTN1,MX$LTA1(P$El1,2)   ; a reading T and first subT wants only
                                          to read it,
        ASSIGN CHTS1,MX$LTA1(P$El1,3)   ; the El1 first copy is taken
        MSAVEVALUE LTA1,P$El1,4,P1
Chakane1  TEST L P1,MX$LTA1(P$El1,2),Vrushta1
        LINK P$El1,FIFO       ; Waiting in the queue of waiting transactions
Vrushta1  ASSIGN Vr,1     ; If T is newer than T, that has blocked El1 first copy
        ADVANCE MX$RAZST(P$Prim1,P2),MX$RAZDEV(P$Prim1,P2)   ; =>
        TRANSFER ,Pat1     ;T returns to its TM
PatPot1  TEST NE P$Nel,1,PatLB1 ; First subT (El1 first copy) is directed to DMP6
        SPLIT 1,Pat1pot,REPLI       ; Sending of confirmation (first copy El1 lock)
        TRANSFER ,PatLB1
Fiksira1  TEST E PBl1,1,RelLock1       ; If T only reads El1
        TEST E V$PRC1,0,ProvCh1     ; Is T the first in the group adjacent reading El1
                                          transactions
ProvCh1  TEST NE MX$LTA1(P$El1,4),P1,RelLock1
        TRANSFER ,Krai1
RelLock1  MSAVEVALUE LTA1,P$El1,1,0   ;Releasing El1 in LM of El1 first copy
        MSAVEVALUE LTA1,P$El1,2,0
        MSAVEVALUE LTA1,P$El1,3,0
        MSAVEVALUE LTA1,P$El1,4,0
        ENTER SEGMRW
        TEST G CH*El1,0,Fik1       ; If there are transactions, waiting of lock of El1 first
                                          copy
        UNLINK P$El1,Provch01,1    ; The first one comes out of user chain and will put
                                          lock of El1
Fik1  LEAVE SEGMRW
        ADVANCE 3,1
        TRANSFER ,Krai1
Provch01 TEST E P$NoEl,1,Provch03 ;
        TEST E P$Bl1,1,Izliza1
        UNLINK E P$El1,Provch01,1,Bl1,1,Izliza1
Krai1  ADVANCE MX$RAZST(P6,P2),MX$RAZDEV(P6,P2)
        TRANSFER ,Krai01                ; SubT, releasing lock of El1 first copy – to the TM
```

264

```
Izliza1  TEST E P$NoEl,1,Zaema3     ; to the LM for El2 first copy lock
       TRANSFER  ,Zaema1                ; to the LM for El1 first copy lock
Pat1pot ADVANCE MX$RAZST(P6,P2),MX$RAZDEV(P6,P2)
       TRANSFER  ,Pat1ob              ; To the uniting subT(El1 first copy) and subT(El1
second copy)
*********************************************************
Pat2   ADVANCE MX$RAZST(P2,P7),MX$RAZDEV(P2,P7)   ; Transfer in the net to
Lock01  ASSIGN 11,V$SPIS2        ; the Lock manager of second copy of El1
       QUEUE P7                     ; Waiting in front of LM of second copy of El1
       SEIZE P7                     ; Processing in Lock manager of second copy of El1
…..
       TEST E CH*11,0,Chaka2            ; Is El1 second copy free,
Zaema2 TEST E MX$LTA2(P$El1,1),0,Chaka2 ; If "YES" then
       MSAVEVALUE LTA2,P$El1,1,P$Bl1     ; the element El1 second copy
       MSAVEVALUE LTA2,P$El1,2,P1        ; from lock table is taken
       MSAVEVALUE LTA2,P$El1,3,P2
       TRANSFER ,PatPt2
…
PatPt2  TEST NE P$Nel,1,PatLB2 ; subT (El1 second copy) is directed to DMP7
       SPLIT 1,PatPot2,REPLI        ; Sending of confirmation (second copy El1 lock)
       TRANSFER ,PatLB2

RelLock2 MSAVEVALUE LTA2,P$El1,1,0  ;Releasing El1 in LM of El1 second copy
…
*********************************************************
Pat3   ADVANCE MX$RAZST(P2,P8),MX$RAZDEV(P2,P8)   ; Transfer in the net to
       ASSIGN NoEl,2                 ; the Lock manager of first copy of El2
Lock03R QUEUE P8                     ; Waiting in front of LM of first copy of El2
       SEIZE P8                     ; Processing in Lock manager of first copy of El2
…..
       TEST E CH*El2,0,Chaka3           ; Is El2 first copy free,
Zaema3 TEST E MX$LTA1(P$El2,1),0,ProvR3 ; If "YES" then
       MSAVEVALUE LTA1,P$El2,1,P$Bl2     ; the element El2 first copy
       MSAVEVALUE LTA1,P$El2,2,P1        ; from lock table is taken
       MSAVEVALUE LTA1,P$El2,3,P2
       TRANSFER ,PatLB3          ; subT (El2 first copy) is directed to DMP8
…
RelLock3 MSAVEVALUE LTA1,P$El2,1,0  ;Releasing El2 in LM of El2 second copy
…
Krai3  ADVANCE MX$RAZST(P8,P2),MX$RAZDEV(P8,P2)
       TRANSFER ,Krai02            ; SubT, releasing lock of El2 first copy – to the TM
*********************************************************
Pat4   ADVANCE MX$RAZST(P2,P9),MX$RAZDEV(P2,P9)   ; Transfer in the net to
       ASSIGN 11,VSPIS2               ; the Lock manager of second copy of El2
Lock03R QUEUE P8                     ; Waiting in front of LM of first copy of El2
       SEIZE P8                     ; Processing in Lock manager of first copy of El2
…..
       TEST E CH*El2,0,Chaka3           ; Is El2 first copy free,
Zaema4 TEST E MX$LTA2(P$El2,1),0,ProvR3 ; If "YES" then
       MSAVEVALUE LTA2,P$El2,1,P$Bl2     ; the element El2 second copy
       MSAVEVALUE LTA2,P$El2,2,P1        ; from lock table is taken
       MSAVEVALUE LTA2,P$El2,3,P2
```

```
     TRANSFER ,PatLB4              ; subT (El2 first copy) is directed to DMP8
…
Fiks4  MSAVEVALUE LTA2,P$El2,1,0  ; Releasing El2 in LM of El2 second copy
…
Krai4  ADVANCE MX$RAZST(P9,P2),MX$RAZDEV(P9,P2)
     TRANSFER ,Krai02           ; SubT, releasing lock of El2 second copy – to the TM
***************************************************
PatLB1  QUEUE FN$Opash1          ; read/write in the local database
        ENTER FN$Opash1          ; with the first replica
        DEPART FN$Opash1         ; of Element1
        ADVANCE 3,1
        LEAVE FN$Opash1
        ASSIGN 5,1
        TRANSFER ,ProvFiks        ; Transmission to the LM of El1 first copy
PatLB2  QUEUE FN$Opash1          ; Record of the second replica of Element1
…
PatLB3  QUEUE FN$Opash1          ; Record of the first replica of Element1
…
        TRANSFER ,LockRW3         ; Transmission to the LM of El2 first copy
PatLB4  QUEUE FN$Opash4          ; Record of the second replica of Element2
…
***************************************************
Krai01   TEST NE PBl1,1,PKrai    ;Assemble the subT, which
        GATHER 2                 ;processed Element1
        ASSEMBLE 2
        TRANSFER ,Krai
Krai02 ……………………………; Assemble the subT, which processed Element2
PKrai   TEST NE P$Nel,1, Krai
        ASSEMBLE 2
Krai    GATHER 2         ;Merging of successfully finished their work
        ASSEMBLE 2       ;(sub)transactions
        TEST NE P$Nel,1, Saber1
        SAVEVALUE ZAVTR2+,1      ; Count committed transactions
        TRANSFER ,Napus          ; accessed two elements
Saber1  SAVEVALUE ZAVTR1+,1      ;accessed only one element
Napus  SAVEVALUE ZAVTR+,1

  DEPART TOTALTIM
  TERMINATE 0
```

Fig. 2. GPSS World model of Distributed 2PL with timestamps

In Fig. 3 we have a summary where the results from throughput capability in relatively minimal, average and maximum regime of intensity of the incoming stream of the transactions of the developed 2PL model and two-version 2PL model are compared. Furthermore, they are almost without deflection from the throughput for system managing database given at [2].
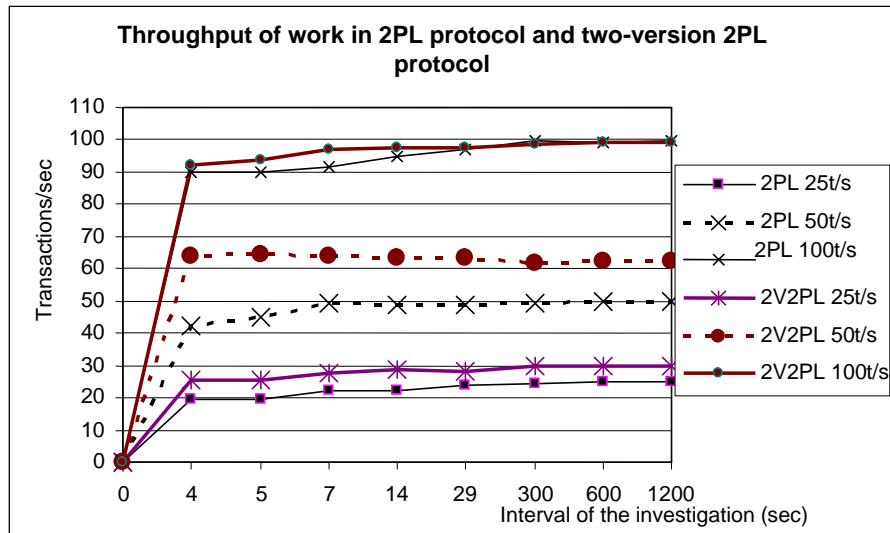
Fig. 3. Throughput of the system

The comparison of the graphs for frequency distribution of response times for distributed 2PL given in fig. 4 with the template one of [2, p.74] shows that the average response time (0.312) and the standard deviation (0.121) are very good results.
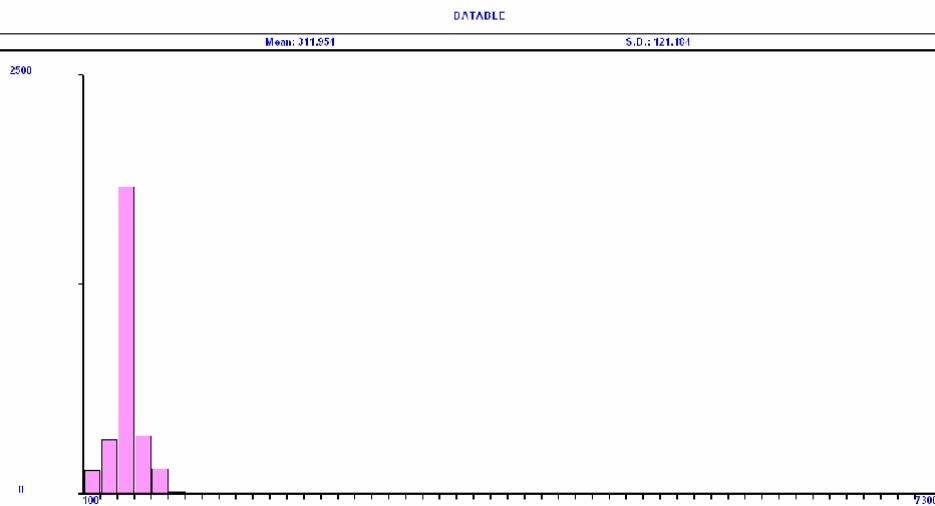
Fig. 4. Frequency distribution of time service

Conclusion

The developed simulation model for investigation of the process processing of transactions in the system with distributed databases allows exclusion of the mutual lockings. The use of the mechanisms for transactions division and submission of certain values and their parameters makes the receiving of results from the execution of transactions in systems with distributed databases possible. The created simulation model describes the real processes with a sufficient accuracy and allows receiving a reliable estimate for the changes of the throughput capability of the system in given parameters of the incoming transaction streams.

References

1. Srinivasa, R., C. Williams, P. Reynolds. Distributed Transaction Processing on an Ordering Network, Technical Report CS-2001-08, February 2001, http://citeseer.ist.psu.edu/cache/papers /cs/26119/srinivasa01distributed.pdf
2. TPC Benchmark^{TM}C, 2006
3. Zhekov, B., S. Vasileva. Opportunities and peculiarities of GPSS World in simulation of concurrency control algorithms for DDBMS. International Scientific Conference Informatics in scientific Knowledge – 2010, Varna Free University "Chernorizets Hrabar" and The Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences, June 24-26, Varna, 2010, pp. 306-318.
4. Бодягин, И. Deadlocks. Что такое взаимоблокировки и как с ними бороться, RSDN Magazine #5, 2003, http://sasynok.narod.ru /index.htm? omvs.htm.
5. Кузнецов, С. Базы данных. Вводный курс. http://www.citforum.ru/database/ advanced_intro/43.shtml, 2008
6. Хасанова, Н. Разработка алгоритмов управления транзакциями, основанных на методе временных меток в распределенных базах данных, 2004, http://science.az/autoreferats/ referat_hasanova_nazli.pdf.

Assoc. Prof. Branimir Kirilov Zhekov, Defence Advanced Research Institute, Phone: +359(0)2 92 26 536, E-mail: bzhekov@yahoo.com
Assist. Prof. PhD Svetlana Zhelyazkova Vasileva, College - Dobrich, University of Shumen, Phone: +359(0)886 491 157, E-mail: svetlanaeli@dobrich.net

CONCEPT OF DATA MANAGEMENT SYSTEMS BASED ON INDIVIDUAL OBJECT PROPERTIES

Vadim Zuravlyov, Eleonora Latisheva, Jurijs Lavendels

__Abstract:__ This paper examines the concept of data management systems based on individual object properties. Such systems operate with objects which can be considered to be physical objects of the real-world and have their unique properties, i.e., the objects have physical properties and/or there are operations to be applied to objects in order to ensure their management. Examples include the autonomous processes of cooperative logistics, cleaning, repair, and maintenance services, smart buildings, monitoring, information, and control systems.

__Key words:__ IS, DMS, DARSIR

1. Introduction

For decades now, the impact of information technology on people's lives has been increasing very rapidly. With every year, one can observe the development of information technologies and the increasing amount of information circulating in information systems. Computer power, once available only to desktop computers, is now available in mobile devices. All this provides new opportunities for developing DMSs (Data Management Systems), as it allows to organize storage and management of local data.

The classic data management theory suggests that one of the most efficient types of data storage and management are centralized DMSs. Development of decentralized DMSs began later, and such topics as multi-agent systems and sensor networks were rapidly spreading during the 1990s. Application of decentralized DMSs proved their effectiveness in various sectors which previously were dominated by the centralized DMSs. These facts suggest a hypothesis that in the range of tasks performed by information systems based on individual object properties, some tasks can be effectively implemented by decentralized DMSs, where the system's object properties (attributes and behavioural rules) can be stored in the object itself. This, in turn, will reduce the total IS construction costs, since there is no longer any need for external data storage and management.

2. Characteristics of a Information System Based on Individual Object Properties

Systems based on individual object properties operate with objects which are considered to be physical objects of the real world, with their unique individual properties (object's physical properties and/or possible

actions with them). Some examples of tasks that are based on object individual properties: autonomous cooperative processes of logistics, cleaning, repairing, and maintenance services, smart buildings, monitoring, information, and control systems.

Mandatory requirements are stated and justified for developing such systems:

- distribution and decentralization of computing resource management;
- support of modern standards for selection of hardware and software;
- support for heterogeneity of physical objects;
- local access to data and their security, privacy, and autonomous operation;
- possibilities to create small solutions and implement them gradually.

Monolithic and decentralized DMSs do not grant the ability to fully comply with the requirements stated, therefore, it is useful to develop a new concept of DMSs based on object individual properties.

3. Selection of technology for new system, based on individual object properties

The name "concept" comes from the Latin word *conceptio* that means "a comprehending, a collection". In this work, the concept of DMS is considered to be classified and summarized views on such system and its processes and to be one of the main steps in the development or modernization of the system. For DMS description, most commonly used data models essentially are abstract, logical definitions of the objects, operators, and other elements that together form an abstract data access machine which the user interacts with. Objects provide modelling of data structures and operators – the data behaviour which forms constructions of abstract programming language that can be used for implementing a wide range of tasks.

Every DMS solution is based on specific technologies. Overview and comparison is given on the common technologies (barcode, RFID, Bluetooth, GSM, GPRS, Wi-Fi, WiMAX, and others) that can be used in systems based on individual object properties.

It is proposed to divide all technologies in three groups, judging by the purpose of using the technology (Table 1).

Table 1. Groups of technologies, by purpose of their use

Nº	Name	Description
I	Technologies for object identification	The main purpose for using the technologies of this group in to identify an object. It is usually a physical object. Discussed technologies are: barcode, RFID (Radio Frequency IDentifikation), NFC (Near Field Communication), etc.
II	Technologies for data transmission	These are technologies, the main purpose of which is to transmit data from one location to another. Discussed technologies are: Wi-Fi, WiMAX, Bluetooth, GPRS (Global System for Mobile Communications), IrDA (Infra-red Data Association), Satellites, etc.
III	Other technologies and sensors	This group covers technologies that are used in conjunction with technologies of Group 1 and Group 2, and are needed to expand the functionality of standard technology. Typically, these are various types of sensors: GPS (Global Positioning System), thermometers, etc.

Technology comparison criteria and their description can be seen in the Table 2.

Table 2. Criteria for comparing technologies

Nº	Name	Description
I	Group of technologies	Three groups of technologies are defined, and they refer to the purpose of use: object identification (Group 1), data transmission (Group 2) and other technologies and sensors (Group 3). One technology can belong to several groups.
II	Working radius	The maximal working radius between two devices.
III	Data transfer rate	How fast data can be transferred from one device to the other. In technologies used for identifying objects, the time required for identifying a single object is indicated.
IV	Working frequency	Most of technologies use a specific frequency. It is important to know them, since each country has its own limitations, and there are technologies use of which is prohibited because of their used frequencies.

V	The need of direct visibility	One of the most important restrictions concerning technologies.
VI	Costs of implementation	Relative costs are given, expressed in relative units, where the first unit means costs of infrastructure, the second unit – costs of terminal part, and the third – costs of tags (for Group 1).
VII	The complexity of technology actuating mechanism	It is the complexity of implementing principles of operation for technology's device. For example, satellite technology is considered to be one of the most sophisticated technologies because satellites should be provided in order to actuate it. The values themselves are defined by technology description and judgements of the field's experts.
VIII	Data protection	For data security and privacy, the level of support provided is defined. Two features are defined: software and hardware.

4. Description of DARSIR concept

The new concept is called Data And Rules Saved In Resource (DARSIR) [1], and it is based on object-oriented approach in accordance to which one should examine the data that are used for system to fulfil its task. DARSIR considers resources to be such data. This covers any living or non-living object that is involved in process of operating information system. Physical presentation of resource acquires the tag: $B = \{S, D\}$, where S is a set of sensors and $D = \{A, N\}$ is the document that for object properties A and rules provides structured storage: $N = \{M, P, F, T\}$ (M – variables, P – procedures, F – functions, and T – trigger—it is a set of operations that is stored within the resource and is run automatically if an event is previously defined). Tag is a part of the concept technology. Within the tag, a link between sensors and document is established. Values received by the sensor can be handled by the rules, as defined in the document. Tag is attached to or built in the resource.

Resources can be divided into Passive Resources or PR (data and attached sensors) and the Active Resources or AR (data with functionality). Active Resources can be divided into two groups: Static (ASR) and Dynamic (ADR). The basic description of ADR is an autonomous operation with the ability to make decision on the site, based on available data. The simplest scenario is shown in Fig 1.
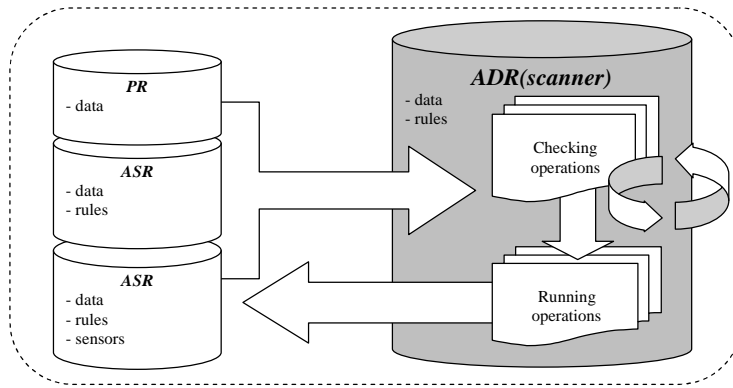
Figure 1. Scenario of single ADR interaction

ADR reads all the data and rules in its memory from Passive Resources (PR) and Active Static Resources that are situated within its reading range. Attributes are received from PR. Rule can also be received from ASR. If ASR rules state that pre-defined activities should be fulfilled, then ADR performs them. If there are rules written in ADR, they are applied, also.

One of the scenarios for ADR interaction is the basic solution for others, when two or more ADRs are situated in a shared area.

If there are two ADRs with the same or similar tasks, they themselves are expected to be working in parallel and there is no interaction between them. Scenario of two ADRs interacting simultaneously is shown in Fig. 2. ADR is working steadily, and it can re-apply its rules to the resources. Rules are designed so that they can be re-run. There can be situations when the resource is within detection area of two ADRs, but the ADRs do not see each other.
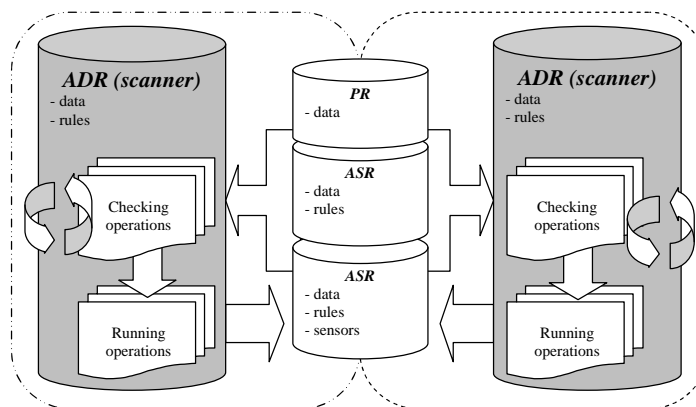
Figure 2. Scenario of two ADRs interacting simultaneously

273

Decentralised data management does not require direct ADR communication. All communications or other forms of interaction take place via the resources (usually PRs and ASRs). ADRs may supplement the resource with data and rules. A similar mechanism exists in the object-oriented approach. It is possible to create complex objects using constructors for composite objects. When these data become unnecessary, they can be deleted from the resource. For example, there is a container for moving boxes of the logistics process (see Fig. 3). The container has its own data (attributes and rules) the initial values of which are defined by the Manufacturer A. When logistics Company B starts using this container, this company adds its part of the data (attributes and new rules). One of the underlying processes of logistics can be implemented by another logistics Company C, which adds its data (attributes and rules) that are necessary for sharing information between members of Company C. When Company C completes its part of the job and returns the container to Company B, the data of Company C (attributes and rules) are no longer relevant and are deleted from the container.
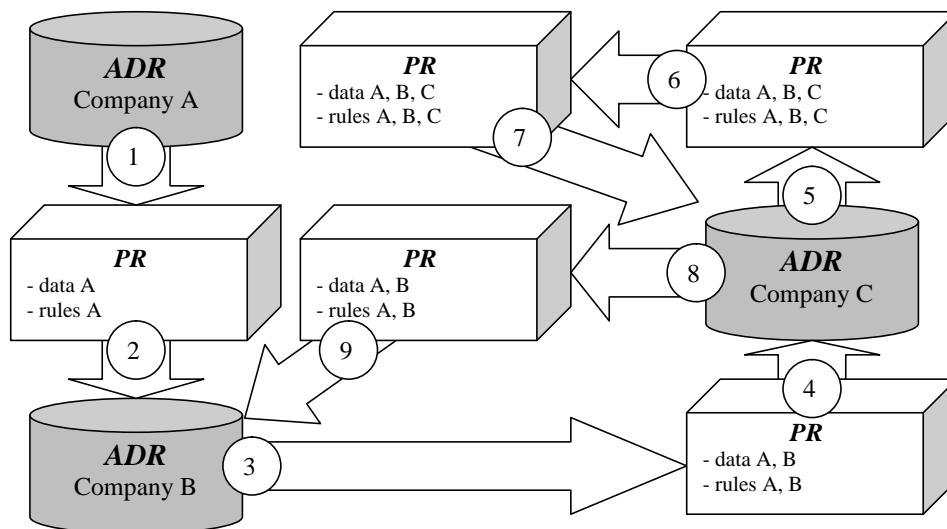
Figure 3. Scenario of consecutive interaction of various ADRs

The resource is described in the DARSIR solution with attributes and rules. One attribute or rule can have several description schemes. One of the ways to streamline working with resources is to classify them, standardize, and divide attributes and rules into groups (see Table 3).

274

Table 3. Attribute and rule groups of DARSIR concept

Name	Description
Resource identifier	Resource has a unique identifier by which to clearly distinguish one resource from a group of resources. There are various methods for identification of physical objects. Some of these methods are suitable for resource identification.
Individual properties	Resources have individual properties. The most common individual properties of resources are physical parameters (length, height, weight, etc.), material properties (plastic, metal, etc.), chemical composition (Calcium, Sodium, etc.), and so on. The way of operating with these properties is defined, and so is the possible application of the fundamental base units, such as length, mass, temperature, etc.
Relations	Resource has additional parameters that link the resource to other objects or places. Usually, it is the owner of the resource (e.g. an individual, organization, trademark, etc.) or resource location (e.g. a city, state, zip code, street address, building, room location, etc.).
Time	This is a part of the measurement system which is used to define the sequence of events, to compare their duration and the interval between events. For humans, the best suited periods are second, minute, hour, day, week, month, year, etc..
History	During the life of physical object, changes are happening. A mechanism is needed for storing history. History group is closely related to the time group. In the history attribute, time and state of the physical object is stored, which was recorded at the time.
Positioning	For resource, location can be defined. In sectors such as logistics, this is one of the most essential attributes. Positioning is divided into two broad categories: out-door and in-door.
Access rights	Because of data privacy and security requirements of DARSIR concept, access rights for resources should be regulated. For object-oriented approach, they are provided with the support of principles of encapsulation. Resources can be influenced by open rules (methods) and attributes. This group provides regulation of access rights.

Classification	A resource is a physical object with its unique individual properties. If there are a lot of resources, it is difficult to apply joint operations to them. One of the options for facilitating the handling of resources is resource classification and categorization. One should provide ability to create complex objects using constructors for composite objects.
Configuration	Resource may contain various other objects. A typical example is a container that contains boxes of goods. If the IS does not provide tags for all the resources, its information can be recorded in the basic resource. Thus, the information can be used in IS, but only through these groups.
Presentation	In the resource groups described above, automation was necessary for resource-related works. User must also provide information on resources that is not related to resource physical properties. Those are different types of data (text, video, audio, etc.) that are characterizing the resource and serving to inform the consumer of the resource about the properties of resource usage. Typical examples of resource presentation: user manual, warranty, advertisement (text, images, audio, video, etc.), and so on.

5. DARSIR concept implementation

For DARSIR concept implementation, two data exchange and storage formats are proposed and implemented:

1. RPML (Resource Physical Markup Language) – an XML-type language, developed specifically for DARSIR concept. The first definition and description of this language was published in [1].

2. JSONR (JavaScript Object Notation for Resource) – JSON-type language, developed specifically for DARSIR concept.

RPML and JSONR data exchange formats can be considered to be DARSIR data exchange formats. Descriptions of application resources for these formats and their individual properties, as well as descriptions of rule implementation are shown in Fig. 4, 5, 6, and 7.

```
{
    "class":"Apple"
    "data":[
    {
    "color":"red",
    "type":"fruits"
    {
}
```

```
<class label = "Apple">
<data>
    <color>red</color>
    <type>fruits</type>
</data>
```

Block — Tag — Value — Tag — Value

a)   b)

Figure 4. Example resource "Apple"
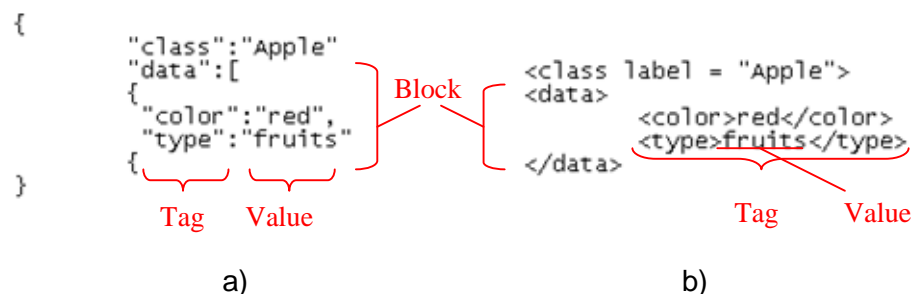
```
{
"epc":"09.34F.16F.1384D32",
"rii":[
{"speed":[
"m":"1","s":"-1","q":"83.3"
],
},
}
```

```
<epc>09.34F.16F.1384D32</epc>
<rii>
        <speed m=1 s=-1>83.3</speed>
</rii>
```

Figure 5. Example presentation of resource individual properties

```
{
"tsk":"switchOn", "params":[{"variable1":"integer"}, {"variable2":"string"}],
"cmd":
[
    {"run":{"func":"switchLightOn", "params":["variable1", "variable2"], "result":"value1"}},
    {"return":{"value1":"bool"}}
]
}
```

Figure 6. Example of JSONR rule

```
<tsk name="switchOn" variable1="integer" variable2="string">
    <run var="variable1,variable2" result="value1">switchLightOn</run>
    <return type="bool">value1</return>
</tsk>
```

Figure 7. Example of RMPL rule

As an auxiliary software for practical creation of DARSIR solution, it is proposed to use a visual interactive simulation. "DARSIR simulation tool" was created which deals with using the DARSIR concept on the basis of RFID technology. The results of proposed approaches are presented in reports [2] and [3].

The developed concept was tested on practical examples [2–7], from which it can be concluded that by using the principles of this concept it is possible to reduce the time needed for creating new solutions.

277

6. Results

The report studied the DMS concept based on object individual properties. Systems based on individual object properties operate with physical objects which are considered to be physical objects of the real-world, with their unique individual properties. Conclusions made during the study:

1. For designing systems based on individual object properties, the following requirements must be met: distribution and decentralization of computational resource management, support of modern standards for hardware and software selection, support for heterogeneity of physical objects, local access to data and their security, confidentiality and autonomy, ability to create small solutions and implement them gradually.

2. Monolithic and decentralized DMSs don't provide full compliance with the requirements stated, therefore, it is useful to develop a new concept of DMS based on object individual properties.

3. The new DARSIR concept is based on object-oriented approaches and considers data to be resources the physical presentation of which is provided with a tag. A tag is a part of the concept technology that stores all object's properties and its possible actions, in a structured way.

5. The resources of DARSIR concept can be divided into Passive Resources or PR (data and attached sensors) and the Active Resources or AR (data with functionality) which, in turn, can be Static (ASR) and Dynamic (ADR) and allow to implement different interaction scenarios.

6. For implementation of DARSIR concept, two data exchange and storage formats are used – RPML and JSONR, which allow to describe data and operations applied to them.

7. Testing of DARSIR concepts on practical examples suggests that using the concept's principles of operation may reduce the time needed for creating new solutions.

References
1. Zuravlyov V. Main Principles of a New Concept of Designing Data Management Systems // Scientific Journal of RTU. 5. series., Datorzinātne. - 30. vol., 2007, pp 38-46.
2. Zuravlyov V., Latiseva E., Karish V. The Theoretical Basis and Practical Usage of the Simulation for Creating IT solutions for Data Management Systems // Proceedings of the international conference on Applied Computer Systems, ISSN 1407-7493, Latvia, Riga, 2011., pp. 138-144.
3. Latiseva E., Zuravlyov V. Various Aspects of Simulation's Usage for Creation of Real Solutions with „DARSIR" Conception as Example // Proceedings of Conference 'Modeling of business, industrial and transport systems 2008;

MBITS 2008', ISBN 978-9984-818-04-7, Latvia, RIGA, 7.-10. May, 2008. - pp 132-135.

4. Zuravlyov V., Matrosov A., Rutko D., Behavior Pattern Simulation of Freelance Marketplace, Workshop on Agents and Multi-agent systems for Enterprise Integration, ZOCO 2012, Salamanca, Espana, March, 2012.

5. Zuravlyov V., Matrosov A., Rutko D., Freelance resource management system optimization, The 2nd International Conference on Computer and Management, CAMAN 2012, Wuhan, China, March, 2012.

6. Latisheva E., Zuravlyov V. Various Aspects of RFID Technology and Their Use in Monitoring and Management of Traffic // Scientific Journal of RTU. 5. series., Datorzinātne. - 34. vol., 2008, pp. 238-245.

7. Latisheva E., Zuravlyov V. Various Aspects of 'Smart Laundry' Task Implementation, Using 'RFID' Technology // Proceedings of papers 'Informatics in the scientific knowledge 2008', Bulgaria, Varna, 26.-28. June, 2008, pp. 150-156.

About the authors

Mg. sc. ing. **Vadim Zuravlyov**, Faculty of Computer Science and Information Technology, Riga Technical University, phone (371) 28346517, email: zuravlov@gmail.com

As. prof. **Eleonora Latisheva**, Faculty of Computer Science and Information Technology, Riga Technical University, phone: (371) 67089576, e-mail: elat@inbox.lv

As. prof. **Jurijs Lavendels**, Faculty of Computer Science and Information Technology, Riga Technical university, phone (371) 67089573, email: jurisl@cs.rtu.lv

RESEARCH OF DATA STORAGE AND MANAGEMENT TECHNOLOGY APPLIED TO DARSIR CONCEPT

Vadim Zuravlyov, Eleonora Latisheva

Abstract: *The paper examines data storage and management technologies that can be used in data management systems based on individual object properties which correspond to DARSIR concept. By the purposes of technology usage, three groups of technologies are defined: object identification technologies, data transmission technologies, and technologies for specific purposes. Based on requirements defined by the DARSIR concept, three groups of comparison parameters were stated for comparing technologies: physical parameters and the external environment, the costs, security and confidentiality. Each technology has defined characteristics, limitations, and specific weight coefficients that can be considered as factors of impact. A formula is determined, and by its calculation it can be concluded that the most suitable technologies for DARSIR concept implementation is 2D barcode and RFID. DARSIR concept with the use of RFID technology has been tested on practical examples "Smart laundry" and "Traffic monitoring and management".*
Key words: *DARSIR, IS, DMS, RFID*

1. Introduction

A range of data management system (DMS) tasks require a variety of independent collaboration elements which are difficult or even impossible to solve by using monolithic system architecture. Examples of tasks that are based on object individual properties:

- Autonomous cooperative processes of logistics;
- Cleaning, repairing, and maintenance services;
- Smart buildings;
- Monitoring, information, and control systems.

Such systems operate with physical objects that have unique individual properties. Let's call them DMSs based on object individual properties. Usually, the most common solution for this kind of systems is Monolithic Data Management Systems (MDMSs) with centralized data storage and management. However, if one needs to provide only limited functionality and bring data handling processes closer to management objects themselves, the MDMS approach becomes uneconomical and it is necessary to create new DMS concept and select data storage and management technologies suitable for it. DMSs with decentralization of data can be carried out more effectively if the object properties (attributes and behaviour rules) are stored in the object itself, and with the use of appropriate technology one can reduce the total IS construction costs.

2. Framework trends and requirements of information systems based on individual object properties

In order to define requirements for systems based on individual object properties, IS framework principles must be defined. Four IS trends from [1] are combined in the Table 1.

<div align="right">Table 1</div>

<div align="center">Basic trends of IS development</div>

N°	Name	Description
I	Distribution and decentralization of computing resource management	With every year, computer systems are becoming smaller and more powerful. Using data communication features (local and wireless networks), it is possible to exchange information between devices. Almost all the necessary tasks users can implement locally, without connecting to large servers.
II	Heterogeneity of IS components	Data are stored in various formats, the handling of data is provided by different types of software and hardware.
III	Distribution of standards	Implementation of standards reduces consequences of Trend II. Users need freedom in choice of hardware and software. It establishes free competition and simplifies IS distribution.
IV	Real-world modelling in ISs	One of the major shortcomings in implementing ISs is the difference between processes in IS and in the real-world. One example is the use of object-oriented models. Such modelling is the principles that were applied to programming and data management. By using it, there is even closer relation between task designing and implementation.

Summarising data from Table 1 and taking into account the range of the studied systems, for systems based on individual object properties the following requirements can be stated (see Table 2).

Table 2

The requirements stated for systems based on individual object properties

N^o	Name	Description
I	Distribution and decentralization of computing resource management	Centralized management can be discarded, if the computing power of local computer systems is sufficient for solution implementation.
II	Using standards	In order to grant IS users free choice of hardware and software, support of modern standards should be provided. For hardware selection process, freedom is especially important. If new solution is implemented, the usage of existing hardware solutions reduces the costs of implementation.
III	Supporting the diversity of physical objects	System must be able to handle large number of objects. Every object has its own individual properties. One object can have unique individual properties that are not suitable for other objects.
IV	Provision of local data access	Acquiring information from sources that are not related to object itself, engages additional resources (data transmission to storage, taking data from storage, providing data exchange interface, organizing data storage space, facilitating data exchange via variety of storages, etc.). If data can be stored in a universal location (for example, within the physical object itself), the costs of data provision can be reduced significantly.
V	Autonomous operation	If every element of DMS operation has access to local data that are necessary for decision-making, it can be said that autonomous operation is ensured.
VI	The economic benefits of implementing small solutions	If the number of physical objects is low, there should be an economic justification for applying this principle of data management. When a small number of object interactions should be provided, large DMS solutions are not economically justified.
VII	Possibility of gradual implementation and distribution of	The requirement VI of small solutions, in turn, requires low implementation costs. Implementation of IS is usually needed to

		implementation costs	automate the manual work, and high implementation costs can reduce the economical advantages of IS implementation.
VIII		Ensuring data confidentiality, integrity, and accessibility	Information security covers all information processes, both physical, as well as electronic, regardless of whether they involve people and technologies, or relationships with partners, customers, or third parties. This applies to all aspects of information and its protection across all information's life-cycle in the organization.

After analysing the most popular monolithic and decentralised DMSs, it was determined that none of them complies to the set requirements for systems based on individual object properties. It became necessary to create a decentralized DMS concept based on object individual properties.

3. DARSIR concept and its corresponding data storage and management technologies

Data And Rules Saved In Resource (DARSIR) [2] concept is based on object-oriented approach in accordance to which one should examine data that are used for system to fulfil its task. DARSIR considers resources to be such data. This covers any living or non-living object that is involved in process of operating information system. The physical presentation of a resource is assigned with a tag that is a set of sensors and documents. The document provides storage for object properties and activities with them (rules). Tag is a part of the concept technology. Within the tag, a link between sensors and document is established. Values received by the sensor can be handled by the rules, as defined in the document. Tag is attached to or built in the resource.

This concept can use a variety of data storage and management technologies:

- Barcode – this technology is used to identify physical objects, there are cases when, along with the identifier, the tag stores additional information about the object (e.g. lot number, manufacturer ID, etc.).

- RFID (Radio Frequency IDentification) – this technology is used for the same purposes as the barcode technology. RFID is usually referred to as "the next generation barcode technology."

- NFC (Near Field Communication) – modification of RFID technology, often used in built-in mobile devices, some of the most typical examples are quick-payments and user identification.

• Local area network – technology that is used for data transmission. For establishing a network, it is necessary to build the infrastructure.

• Bluetooth, IrDA (Infra-red Data Association) – data transmission technologies, typically used for contacting mobile devices.

• GSM (Global System for Mobile Communications), GPRS – mobile communication standard and mobile data transmission service, used for data transmission, now widely accessible and used in locations where other transmission standards are not available.

• Wi-Fi, WiMAX – wireless data transmission technology, nowadays its infrastructure is widely used.

• Satellites – one of the most popular examples of this technology is GPS (Global Positioning System), that is suitable for positioning. Similar techniques that work on the same principles are Galileo, KOMPASS, GLONASS, and others.

The concept of decentralised data management system based on individual object properties should meet requirements defined in Table 2. By the purposes of technology use, it is proposed to divide all technologies into three groups (see Table 3), and the parameters for comparing technologies are described in Table 4.

Table 3

Groups of technologies, by purpose of their use

N°	Name	Description
I	Technologies for object identification	The main purpose for using the technologies of this group in to identify an object. It is usually a physical object. Technologies discussed are: barcode, RFID, NFC, etc.
II	Technologies for data transmission	These are technologies, the main purpose of which is to transmit data from one location to another. Discussed technologies are: Wi-Fi, WiMAX, Bluetooth, GPRS, IrDA, Satellites, etc.
III	Other technologies and sensors	This group covers technologies that are used in conjunction with technologies of Group 1 and Group 2, and are needed to expand the functionality of standard technology. Typically, these are various types of sensors: GPS, thermometers, etc.

Table 4

Criteria for comparing technologies

N°	Name	Description
I	Group of technologies	Three groups of technologies are defined, and they refer to the purpose of use: object identification (Group 1), data transmission (Group 2) and other technologies and sensors (Group 3). One technology can belong to several groups.
II	Working radius	The maximal working radius between two devices.
III	Data transfer rate	How fast data can be transferred from one device to the other. In technologies used for identifying objects, the time required for identifying a single object is indicated.
IV	Working frequency	Most of technologies use a specific frequency. It is important to know them, since each country has its own limitations, and there are technologies the use of which is prohibited because they use specific frequencies.
V	The need of direct visibility	One of the most important restrictions concerning technologies.
VI	Costs of implementation	Relative costs are given, expressed in relative units, where the first unit means costs of infrastructure, the second unit – costs of terminal part, and the third – costs of tags (for Group 1). Unit costs are within the range of 0 to 9 (the higher, the more expensive).
VII	The complexity of technology actuating mechanism	It is the complexity of implementing principles of operation for technology's device. For example, satellite technology is considered to be one of the most sophisticated technologies because satellites should be provided in order to actuate it. The values themselves are defined by technology description and judgements of the field's experts.
VIII	Data protection	The level of support provided is defined for data security and privacy. Two features are defined: software (SW), and hardware (HW).

For each technology, characteristics and limitations are defined, thus allowing to create table of comparison for technological parameters (see Table 5).

Table 5

Table of comparison for technology parameters (1 – group of technologies; 2 – working radius (km); 3 – data transfer rate (MHz, or obj./s); 4 – working frequency; 5 – the need of direct visibility (no, preferably, yes); 6 – implementation costs (infrastructure / terminal / tag); 7 – complexity of technology actuating mechanism; 8 – data protection)

Parameter \ Technology	1	2	3	4	5	6	7	8
Barcode (Code128)	1	0.01	0.3–0.5 obj./s	n/a	yes	3/1/0	no	SW
Barcode (DataMatrix)	1	0.002	0.2–0.4 obj./s	n/a	yes	3/2/0	no	SW
Barcode (QR Code)	1	0.002	0.2–0.4 obj./s	n/a	yes	3/2/0	no	SW
RFID (passive)	1	0.0001–0.009	1–200 obj./s	LF, HF	pref.	4/2/1	no	SW, HW
RFID (semi active)	1	0.001–0.035	1–200 obj./s	UHF, MF	pref.	4/2/3	no	SW, HW
RFID (active)	1	0.035–0.1	1–200 obj./s	UHF, MF	pref.	4/2/5	mid	SW, HW
NFC	1, 2	0.0015	424 Kb/s	13,56 MHz (HF)	pref.	4/1/1	no	SW, HW
LAN	2	0.1	1 Gbit/s	10 MHz	no	5/3/-	mid	SW, HW
Bluetooth	1, 2	0.1	2 Mb/s	2.402–2.483 MHz	pref.	2/2/4	no	SW
IrDA	1, 2	0.01	4 Mb/s	32–40 KHz	yes	2/2/4	no	SW
GSM	1, 2	35	43.3 Kb/s	380–1900 MHz	no	7/3/5	yes	SW, HW
GPRS	2	35	80 Kb/s	380–1900 MHz	no	7/3/5	yes	SW, HW
Wi-Fi	1, 2	0.1	600 Mb/s	2.5–5 GHz	pref.	4/3/6	yes	SW, HW
WiMAX	1, 2	50	70 Mb/s	2–66 MHz	pref.	5/3/7	yes	SW, HW
GPS	3	global	n/a	1227–1575 MHz	yes	9/2/2	yes	SW, HW
Sensors	3	0.001	n/a	n/a	yes	2/2/2	no	n/a

Table analysis suggests that:

• Single technology can correspond to only one group of technologies or to Group 1 and Group 2 simultaneously, and each technology has different technology parameter values.

286

• None of the described technologies solves the issue of data confidentiality, integrity, and accessibility. This requirement should be supported on the software level.

• Only Group 1 and 2 technologies can be used as basic technologies, while practically implementing DARSIR concept. The technologies of Group 3 can only be used as additional source of information.

4. Selection of technology

DARSIR concept can be used with various technologies. A simple and reliable method is needed, by which one can quickly select the optimal technology, based on solution's needs. During the analysis and development of solution, new requirements can arise, this method should quickly assess the impact on the technology selection. The basic criteria are examined that can influence this selection. Based on these criteria, the technology selection calculation method is defined. For all criteria, values of equal significance are defined, and the technology is selected that by default is the most appropriate for DARSIR concept.

It is necessary to classify the impact criteria and to define sets of criteria:

• Physical parameters and the external environment – one needs to define parameters that affect the mandatory implementation of the solution, i.e.: tag parameters (size, storage volume, ability to overwrite, etc.), scanner parameters (scanning rate, distance to the object), environmental conditions (temperature, humidity, etc.). etc.

• Costs – the total costs of running and maintaining the solution.

• Security and privacy – implementing data protection methods of various kinds or combination of such methods.

Specific technology is being chosen by assessing the need for IT solution. Usually, a compromise is found between options and costs. Technology criteria are defined, and every technology has its own parameter of technology type. For every type of technology parameter, a value is defined in the range from 0 to 1, the higher the parameter value, the better this parameter suits for DARSIR concept solution. For example, tag data storage volume, linear barcode technology. Its volume is 10 bytes and the value is 0.1. The volume for two-dimensional barcodes is 100 bytes and the value is 0.6. For semi-active RFID tags, the volume is 1,000,000 bytes and the value is 0.8.

Another weight coefficient for technology parameters is defined, the permissible value range of which is 0 to 1. For example, when barcode

technology is already used for a specific solution, the weight coefficient of technology type is 1. In cases when there is a need to calculate two-dimensional barcodes, the weight coefficient is 0.5, but RFID technology requires new costs, so the value is 0.1.

For each technology, characteristics and limitations are defined. There are cases when essential technology limitations prevent from using certain technology. If the lowest possible value is set for an essential technology type parameter, then these limits are calculated during technology selection. For example, the tag data storage volume in the previous case. Let's suppose the IT solution requires 90 bytes of data volume, which corresponds to the value of 0.5. For linear barcodes, this value is 0.1, for two-dimensional barcodes it is 0.6, and for semi-active RFID tags it is 0.8. This means that all technologies correspond to the task, with the exception of linear barcodes.

Criteria Coefficient of Used Technology Combination (CCUTC) is calculated according to formula (Formula 1):

$$CCUTC = \sum_{i=1}^{n} a_i h(c_i, t_i) \tag{1}$$

where n is the number of technologies, a_i is the weight coefficient of technology type, c_i is the type parameter of ith technology, t_i is ith is the threshold for acceptable values of ith technology type parameter, and h is the threshold function that acquires the value of $+\bullet$, when c_i does not reach the value of t_i:

$$h(c,t) = \begin{cases} c, & ja \ c \geq t \\ + \infty, & ja \ c < t. \end{cases} \tag{2}$$

Coefficients of important parameters are called Weight Coefficients (WC), and their values can be seen in Tables 6, 7, and 8.

Table 6
WC of physical parameters data volume and external environmental technologies (1 – data amount of tags, 2 – overwriting option, 3 – tag size; 4 – tag weight; 5 – atypical environmental conditions 6 – scanner reading rate; 7 – scanning distance; 8 – direct visibility of the scanning mode; 9 – scanner's multiple tag reading, 10 – technology recognition, distribution, and availability)

WC \ Technology	1	2	3	4	5	6	7	8	9	10
Barcode (Code128)	0.2	0.1	0.9	0.9	0.9	0.4	0.3	0.1	0.1	0.8
Barcode (DataMatrix)	0.6	0.2	0.9	0.9	0.9	0.5	0.2	0.1	0.4	0.8
Barcode (QR Code)	0.6	0.2	0.9	0.9	0.9	0.5	0.2	0.1	0.4	0.8
Bluetooth	0.9	0.9	0.3	0.3	0.1	0.5	0.6	0.7	0.2	0.6
GSM	0.8	0.6	0.3	0.3	0.1	0.3	0.8	0.7	0.7	0.8
IrDA	0.9	0.9	0.3	0.3	0.1	0.5	0.6	0.3	0.2	0.5
NFC	0.6	0.5	0.9	0.9	0.7	0.7	0.6	0.7	0.4	0.6
RFID (active)	0.9	0.9	0.2	0.5	0.1	0.5	0.9	0.8	0.7	0.4
RFID (passive)	0.6	0.1	0.9	0.9	0.9	0.7	0.6	0.7	0.9	0.8
RFID (semi active)	0.8	0.7	0.8	0.8	0.2	0.7	0.8	0.7	0.9	0.6
Wi-Fi	0.9	0.7	0.2	0.2	0.1	0.3	0.7	0.7	0.6	0.9
WiMAX	0.9	0.9	0.2	0.2	0.1	0.2	0.9	0.7	0.6	0.5

Table 7
WC of technology costs (1 – costs of technology implementation; 2 – costs of technology maintenance; 3 – scanner costs; 4 – tag costs; 5 – tag's life-cycle)

WC \ Technology	1	2	3	4	5
Barcode (Code128)	0.7	0.8	0.6	0.9	0.9
Barcode (DataMatrix)	0.7	0.8	0.7	0.9	0.9
Barcode (QR Code)	0.7	0.8	0.7	0.9	0.9
RFID (passive)	0.6	0.8	0.6	0.8	0.9
RFID (semi active)	0.5	0.7	0.6	0.7	0.7
RFID (active)	0.4	0.5	0.5	0.6	0.6
NFC	0.7	0.8	0.6	0.8	0.9
Bluetooth	0.5	0.5	0.6	0.6	0.4
IrDA	0.4	0.5	0.6	0.6	0.4
GSM	0.3	0.4	0.4	0.6	0.3
Wi-Fi	0.5	0.5	0.4	0.5	0.3
WiMAX	0.4	0.4	0.3	0.4	0.2

WC of data security and privacy (1 – built-in error detection features; 2 – built-in data recovery features; 3 – built-in features against unauthorized data use; 4 – identification methods; 5 – scanning accuracy)

WC Technology	1	2	3	4	5
Barcode (Code128)	0.4	0.1	0.1	0.9	0.2
Barcode (DataMatrix)	0.7	0.7	0.7	0.4	0.6
Barcode (QR Code)	0.7	0.7	0.7	0.4	0.6
RFID (passive)	0.4	0.1	0.5	0.9	0.7
RFID (semi active)	0.5	0.4	0.6	0.9	0.8
RFID (active)	0.6	0.6	0.7	0.9	0.8
NFC	0.4	0.1	0.5	0.9	0.7
Bluetooth	0.5	0.4	0.8	0.7	0.7
IrDA	0.5	0.4	0.8	0.7	0.7
GSM	0.5	0.3	0.7	0.6	0.8
Wi-Fi	0.5	0.2	0.8	0.7	0.7
WiMAX	0.5	0.2	0.8	0.7	0.7

After performing CCUTC calculations, using WC values of equal weight, it is concluded that the most suitable technologies for DARSIR concept are 2D barcode and RFID technologies. However, RFID technology has broader range of functionality options, compared to 2D barcode technology. For example, for barcodes, it is difficult to implement data overwriting capabilities, a new tag should be printed over the old one.

5. DARSIR concept test on a typical task of "Smart Laundry"

"Smart laundry" is set of requirements for cleaning clothes in an intelligent and automated laundry service. Laundry services include washing, drying, ironing, and professional fabric care. The main requirement of this system is to minimize the human error that might damage laundry during its washing process. The technical requirement ensures the maximum fabric processing allowed and care appropriate to its maintenance instructions. The task of "Smart Laundry" is described in report [3].

The basic resource is laundry, the cleaning service of which must be provided. In the laundry, there is a built-in tag which stores information about fabric parameters (e.g. type, weight, colour, etc.), the steps necessary (washing, bleaching, drying, ironing, and professional fabric

care), and rules that apply to the laundry and are not a part of standard washing cycle. Every resource of the service (service providing equipment and detergents) also has a tag with description.

The basic operation resource is the equipment for provision of service, it is triggered when laundry is placed into a specific machine. Let's assume there is a machine (washing machine) with laundry inside. The machine receives attributes and rules from the laundry, analyses this information, and runs the rules of equipment. After making decisions, a laundry-rule check is performed to test if there are no breaches concerning service parameters selected. Then, detergent selection is made and rules are applied according to the selected detergents. The optimum operation mode is selected and laundry processing begins. When there are restrictions incorporated, if equipment can detect a way to correct the situation and run the operation-cycle, it selects the optimum mode, otherwise the laundry owner or the employee of laundry service is offered the available options for further steps, depending on internal instructions of the service.

A typical algorithm for laundry analysis and selection of the most appropriate operating mode:

1. To check if all the resources located in the washing machine can be washed (parameter Washable), if not, an error message appears: "Not all items in the washing machine are washable."

2. To check if the total weight of all the resources located in the washing machine (parameter Weight) does not exceed the maximum weight allowed in this machine. If weight is exceeded, an error message appears: "Washing machine is overloaded."

3. To check if materials of these resources can be washed together. If not, an error message appears: "Those materials shouldn't be washed together."

4. To select the washing temperature.

5. To select the rotation rate.

6. To run the washing process.

7. The washed resources are shown on the history report of what was washed.

The next step is to select technology the basis of which will be implementing the task "Smart Laundry." For fulfilling this task, it is recommended to choose RFID technology since the barcode technology has significant limitations: barcode tag is easy to damage, tag visibility is required for data reading process, it is not possible to modify data on the tag, etc.

"DARSIR simulation tool" has three types of objects:

• Laundry – a Passive Resource, by which the basic actions can be applied to the laundry service. Resource has recorded attributes that describe the washing properties (such as colour, weight, material, etc.).

• Detergents – an Active Static Resource, the means that are necessary during laundry service operating cycle. This type of resource has recorded internal rules that describe permissible and prohibited activities with detergents. For example, bleach can be used only for white linen.

• Washing machine – machine for providing service, an Active Dynamic Resource that is provided for washing machine. For this resource, rules are attached based on the washing machine operation algorithm.

For creating solution, there is "DARSIR simulation tool" [4]. Smart washing solution "DARSIR simulation tool" is shown in Figure 1b. The solution created by simulation tool is transferred to mobile device application, as shown in Figure 1b.
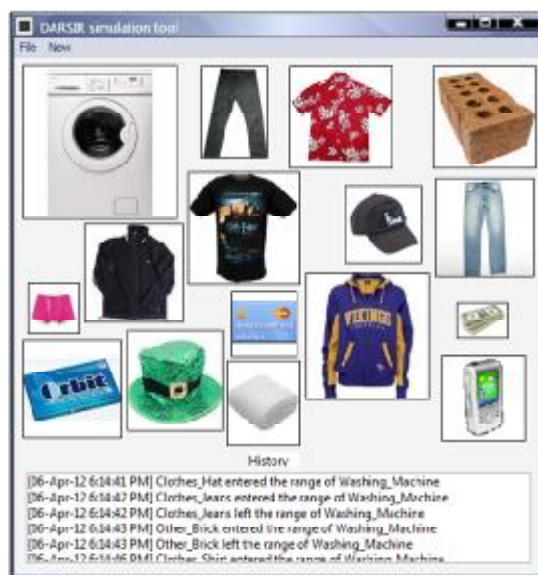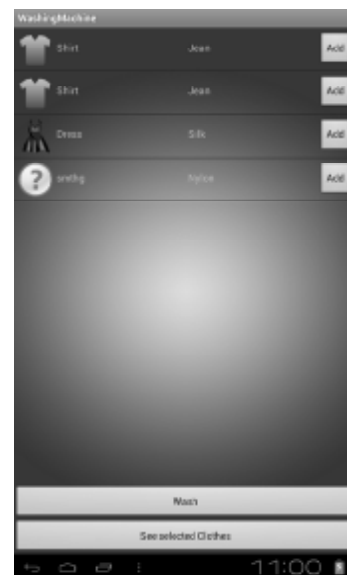
a) Simulation tool   b) Mobile device application

Figure 1. Smart laundry solution

6. Results

The report studied the DMS concept based on object individual properties. Systems based on individual object properties operate with physical objects which are considered to be physical objects of the real-

world, with their unique individual properties. Conclusions made during the study:

1. There is a range of DMS tasks that require a variety of independent collaboration elements which are difficult or even impossible to solve by using a monolithic system architecture. Such systems (systems based on individual object properties) operate with physical objects that have unique individual properties.

2. For designing systems based on individual object properties, the following requirements must be considered: distribution and decentralization of computational resource management, support of modern standards for hardware and software selection, support for heterogeneity of physical objects, local access to data and their security, confidentiality and autonomy, ability to create small solutions and implement them gradually.

3. After analysing the most popular monolithic and decentralised DMSs, it was determined that none of them complies to the set requirements for systems based on individual object properties. It became necessary to develop a decentralized DMS concept based on object individual properties.

4. The new DARSIR concept is based on object-oriented approach and considers data to be resources the physical presentation of which is provided with a tag. A tag is a part of the concept technology that stores all object's properties and its possible actions, in a structured way.

5. With DARSIR concept, a variety of data storage and management technologies can be used, and after comparing them the RFID technology was chosen since it satisfied the most of criteria set forth.

6. DARSIR concept test, using a typical task of "Smart Laundry," allows to make conclusions about the usefulness of DARSIR concept and the selected RFID technology for this kind of tasks.

References

1. Simon, A. R. Strategic Database Technology: Managament for The Year 2000. Morgan Kaufmann Publishers Inc., USA, San Francisco, 1995, p. 479.
2. Zuravlyov, V. Main Principles of a New Concept of Designing Data Management Systems // Scientific Journal of RTU. 5. series., Datorzinātne. - 30. vol., 2007, pp 38-46.
3. Latiseva, E., Zuravlyov, V. Various Aspects of 'Smart Laundry' Task Implementation, Using 'RFID' Technology // Proceedings of papers 'Informatics in the scientific knowledge 2008', Bulgaria, Varna, 26.-28. June, 2008, pp. 150-156.
4. Latiseva, E., Zuravlyov, V. Various Aspects of Simulation's Usage for Creation of Real Solutions with „DARSIR" Conception as Example // Proceedings of

Conference 'Modeling of business, industrial and transport systems 2008; MBITS 2008', ISBN 978-9984-818-04-7, Latvia, RIGA, 7.-10. May, 2008, pp 132-135.

About the authors

Mg. sc. ing. **Vadim Zuravlyov**, Faculty of Computer Science and Information Technology, Riga Technical University, phone (371) 28346517, email: zuravlov@gmail.com

As. prof. **Eleonora Latisheva**, Faculty of Computer Science and Information Technology, Riga Technical University, phone: (371) 67089576, e-mail: elat@inbox.lv

INFORMATICS IN EDUCATION

WIKIS IN E-LEARNING

Teodora Bakardjieva, Antonina Ivanova

Abstract: In this paper educational benefits of wikis are discussed. Varna Free University incorporated a wiki module in its e-Learning system, which encouraged both tutors and students to harness their collective intelligence in order to achieve their common educational goals and the exploration of new and effective uses of the Wiki tool is also presented. Two most important factors to consider when implementing a Wiki are how to integrate the Wiki project and how to assess student learning and participation in the Wiki project. Students are quick to realize when a project is an add-on requirement and perceive this as extra busy work. The teacher must provide a clear assessment procedure that will be as objective as possible. For this purpose student self-assessments can be used and this procedure could also include more specific requirements concerning the number and length of each post, or could include a note grade instead of pass-fail based on content and form.

Key words: web 2.0, e-Learning, wiki, knowledge management

Introduction

Nowadays we're witnessing exploration of the potential of wikis, blogs, media-sharing services and other social software - which, although not designed specifically for e-learning, can be used to empower students and create exciting new learning opportunities.

Chatti et al. [6] discuss the potential use of social software in learning environments. Open blogs and cloud platforms such as Facebook have great educational potential [12]. Setting up an e-learning system is very easy now. Almost anyone can now establish an online learning community using open source learning tools that comprise Web 2.0 features. That's why it is now possible for any organization to afford personalized online courses with a learning management system having advanced features to support mutual communication and collaboration. Production and delivery of e-learning programs is far easier with the arrival of Web 2.0. Discussion forums, wikis, blogs and podcasts are just a beginning in the field of online learning.

Wiki applications facilitate collaborative editing supported by revision mechanisms that allow the monitoring of changes. Wiki technology can be used as a community platform but also as a personal

authoring environment. Wiki was developed in 1994 by Ward Cunningham. Wiki comes from the Hawaiian word "wiki-wiki" meaning fast. "WikiWikiWeb" was created in 1995 by Ward Cunningham as an online manual for software programmers to share knowledge [17]. Jimmy Wales built on this idea and created Wikipedia, and now everybody is familiar with Wikipedia, which is itself a Wiki in the form of an online encyclopedia that can be edited by any user. Educators are now experimenting with using Wikis in pedagogically sound ways. Each user has the ability to modify any part of the Wiki space, analogous to a mini-website. Users create new nodes in the hierarchy each time that they want to elaborate, change or add content. Using Wikis can allow for a numerous opportunities for collaboration between students, but students do not have to be in the same physical location to meet with each other. These kinds of programs "allow for cooperation between the instructor and students or among students by using different formats of social interaction" [7].

Evaluating the quality of contributions in such collaborative authoring environments is a challenging task. However, based on the "wisdom of the crowd" principle, one collects and aggregates enough data until there is a consistently reliable answer. Oren et al. [15] acknowledge that wikis are successful for information collection, but point out that they do not fully satisfy the requirements of PKM. A semantic wiki allows users to make formal descriptions of resources by annotating the pages that represent those resources. Whereas a regular wiki enables users to describe resources in natural language, a semantic wiki allows users to additionally describe resources in formal language. Semantic wikis augment ordinary wikis by using the metadata annotations, and thus may offer better information retrieval and knowledge reuse.

Wikis enable users to collaboratively create and edit web content directly, using a web browser. In other words, a wiki is a collaborative web site whose content can be edited by anyone visiting the site, allowing them to easily create and edit web pages [4]. Wikis can serve as a source of information and knowledge, as well as a tool for collaborative authoring. Wikis allow visitors to engage in dialog and share information among participants in group projects, or to engage in learning with each other by using wikis as a collaborative environment in which to construct their knowledge [3].

As defined in Leuf and Cunningham [11], the proper term "Wiki" is used to refer to the essential concept rather than to any particular implementation, the latter being called simply a "wiki". From a technical standpoint, the Wiki concept rests on the World Wide Web, and the underlying HTTP protocol defines how the client-server communications

296

occur. At the functional level, the essence of Wiki can be summarized as follows:

- a wiki invites any and all users to edit any page or to create new pages within the wiki site, using only a simple web browser without anyadditional add-ons;
- wiki encourages meaningful topic associations between pagesby making the creation of page links almost intuitively easy;
- rather than serving as a carefully crafted site for casual visitors,a wiki seeks to involve the visitor in an ongoing process of creation andcollaboration that constantly changes the web site content;
- semantic Wikis extend wikis with formal annotations describing the content and create views;
- semantic Wikis introduce background knowledge;
- semantic Wikis for PKM – formal structure gives automated support and flexibility of wiki gives people freedom.

Background

Wiki modifications are easy because the processes of reading and editing are both quite simple. In essence, a wiki is a simplification of the process of creating HTML web pages. Simply clicking an "edit this page" link allows instant revisions [10]. Wikis are editable through a browser, and the editing interface is generally simple and easy to use.

Wikis provide a mechanism to record every change that occurs over time as a document is revised. Each time a person makes changes to a wiki page, that revision of the content becomes the current version, and an older version is stored. Versions of the document can be compared side-by-side, and edits can be "rolled back" if necessary. This means that it is possible to revert a page (if necessary) to any of its previous states.

Further, the administrator of the site has control over access, determining which portions are user-editable. Some wikis restrict editing access, allowing only registered members to edit page content, although anyone may view it. Others allow completely unrestricted access, allowing anyone to both edit and view content [13].

Many wiki systems are adding functionalities such as web-based spreadsheets, calendars, documents, photo galleries, private workspaces, hierarchical organization, WYSIWYG (what you see is what you get) web editing, importing Word or Excel files, and even integration with centralized content management systems [10]. WikiMatrix [19] provides a tool to compare the features of various popular wiki engines.

Educational benefits of wikis revolve around the fact that they offer an online space for easy interaction and collaboration. Both teachers and students can easily create web pages using wikis without prior knowledge or skill in web development or programming, eliminating the extra time necessary to develop these skills. A wiki offers the ability to interact with evolving text over time as well, allowing teachers and learners to see assignments as they are drafted, rather than commenting only on the final draft. Considering the complications of scheduling after-hours meetings for students, a wiki can also be extremely useful for communication within groups. Further, as more organizations adopt wikis for internal and external collaboration and information dissemination, interacting with them at the educational level builds important work skills.

Observations and Discussions

Varna Free University incorporated a wiki module in its e-Learning system (Fig 1), which encouraged both tutors and students to harness their collective intelligence in order to achieve their common educational goals and the exploration of new and effective uses of the Wiki tool is also presented.

Moodle has the most transparent and easiest navigation especially for a generation of students well trained in text editing in programs such as Microsoft Word (Fig. 2, Fig. 3, Fig. 4, Fig. 5).

Fig. 1. VFU E-learning System
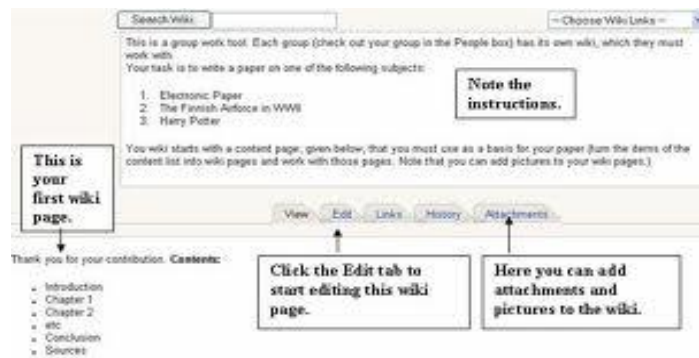
Fig. 2. Wiki HTML Editor

Fig. 3. Wiki in Moodle

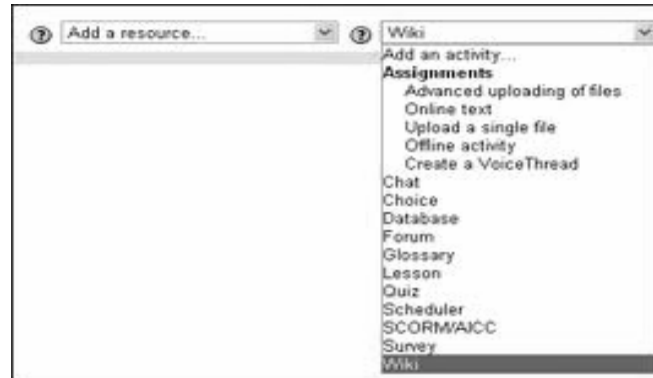Fig. 4. Adding images to a wiki in Moodle

Fig. 5. Moodle Wiki Editing

Wikis find application in the educational process and in the university management system as well. Placement centers use wiki pages to store and organize content for job postings and career development. Wikis act as a sounding board so that students can voice opinions about university policies.

The free-form, collaborative nature of wikis makes them easy to apply in creative ways. Any sort of group process can be facilitated using a wiki. Creating a wiki for group lecture notes after a lecture gives students a chance to combine their notes. Those that missed information get it from their peers. The group decides what information is critical and gives it proper emphasis. Group lecture notes are done.

The most straightforward use of a wiki is as a tool for group collaboration for creating group projects. A teacher assigning a group project gives students a place to work by creating a wiki with the group mode enabled. This gives each group their own space to record research, to develop outlines and to create the final product. The teacher creates a submission date on which to turn off editing capabilities for students so that he/she can grade the final projects. Afterwards, the teacher enables visible groups so that everyone can see each other's work. Also, a teacher develops a wiki for a student group and urges people to submit ideas around a brainstorming topic. People can add ideas as they occur and link to other pages for elaboration.

A teacher might assign students the task of contributing to another wiki on the Web, on any study topic, perhaps by assigning students to groups and challenging them to collaboratively create an article they would feel confident posting to a public-information space. Students use the course wiki to create drafts of the article they will eventually publish to the community at the end of the semester.

This type of assignment has a number of benefits:

• It gives students additional motivation to do their best, since they know their work will be viewed and critiqued by the public instead of just by their instructor.

• It can act as a summarizing activity for an entire semester's work.

• Students will know their work will be used by other people, not just graded and filed away.

An inquiry was made among 257 MBA students and as a result they had mainly positive reactions to the Wiki projects. Some students commented that Wiki was a useful tool and a good way to put a summary of the lecture content together in a way that all students could benefit. Others mentioned that it allowed them to carry on dialogue with varying view-points that offered a more holistic learning experience. There was dissatisfaction about the fact that it is hard to grade participation because oftenpeople post the same things. Pointing out the pedagogical benefits from the project we have to stress on the assessment. As seen in the survey results above, this is one area in which students were the least satisfied. Students were assessed on their participation in the Wiki. Any user can see who has made a contribution, the date and time of each contribution.

Table 1: Student Survey of Wiki Projects
(MBA 257 – 60 % response rate)

Response	Ease of use	Enhanced learning	Useful for exam prep	Fair assessment of students' efforts
1 – strongly agree	15	13	25	22
2- agree	65	24	27	40
3 – neutral	13	59	30	17
4 – disagree	7	4	18	21
5 – strongly disagree	0	0	0	0

Another difficulty in integrating Wikis successfully comes from the switch to a student-centered approach. Using student-created Wikis as a major content source shifts the creation and ownership of knowledge base from the teacher to the student. The role of student in this Wiki projectis that of primary content producer. The teacher's role changes to one of facilitating and correcting errors. As mentioned above, Wikis are quite straightforward, and Moodle has Wiki interface that resembles common word processing programs, so students may find it easy to use.

Conclusions

Using Wikis in an LMS such as Moodle can be a useful teaching tool and can support a student-centered learning experience. Maybe the two most important factors to consider when implementing a Wiki are how to integrate the Wiki project and how to assess student learning and participation in the Wiki project. Students are quick to realize when a project is an add-on requirement and perceive this as extra busy work. The teacher must provide a clear assessment procedure that will be as objective as possible. For this purpose student self-assessments can be used and this procedure could also include more specific requirements concerning the number and length of each post, or could include a note grade instead of pass-fail based on content and form.

By carefully designing the implementation, integration, and evaluation of a Wiki, a new, enjoyable collaborative space can be created which allows for much more efficient, asynchronous, and evaluated interaction between groups of students and teachers. Also, the flexibility of this medium allows for extremely varied adaptations in an extraordinarily large range of educational contexts.

Some educational uses of wikis can be outlined:

Ø Students use a wiki to develop research projects, with the wiki serving as ongoing documentation of their work.

Ø Students add summaries of their thoughts from the prescribed readings, building a collaborative annotated bibliography on a wiki.Wiki is used for publishing course resources like syllabi and handouts, and students comment on these directly for all to see.

Ø Teachers use wikis as a knowledge base, enabling them to share reflections and thoughts regarding teaching practices, and allowing for versioning and documentation.

Ø Wikis map concepts. They are useful for brainstorming, and editing a given wiki topic can produce a linked network of resources.

Ø Wiki is used as a presentation tool in place of conventional software, and students are able to directly comment on and revise the presentation content.

Ø Wikis are tools for group authoring. Often group members collaborate on a document by emailing to each member of the group a file that each person edits on their computer, and some attempt is then made to coordinate the edits so that everyone's work is equally represented in a single, central wiki page.

Wikis usage in an e-Leaning system can be a powerful teaching tool that enhances and increases collaboration outside of class.

References

1. Bates, T. (2011) E-Learning Outlook for 2011. http://www.tonybates.ca/2011/ 01/16/e-learning-outlook-for-2011/

2. Bates, Tony. (2010) Blackboard acquires Elluminate and Wimba: the end of LMSs? http://www.tonybates.ca/2010/07/11/blackboard-acquires-elluminate-and-wimba-the-end-of-lmss/

3. Boulos, M. N. K., Maramba, I. & Wheeler, S. (2006) Wikis, blogs andpodcasts: A new generation of web-based tools for virtual collaborativeclinical practice and education, BMC Medical Education, 6(41). http://www. biomedcentral. com/content/pdf/1472-6920-6-41.pdf (Last visited May 2012)

4. Chao, J. (2007) Student project collaboration using Wikis. Proceedings of the 20th Conference on Software Engineering Education and Training (CSEE&T2007), Dublin,Ireland: July 3-5.

5. Chatti, M. A., Jarke, M. and Frosch-Wilke, D. (2007) The Future of e-Learning: a Shift to Knowledge Networking and Social Software. International Journal of Knowledge and Learning, IJKL, 3 (4), p. 404-420

6. Chatti, M. A., Srirama, S., Kensche, D. & Cao, Y. (2006) Mobile web services for collaborative learning. Proceedings of the 4th International Workshop on Wireless, Mobile and Ubiquitous Technologies in Education (WMUTE 2006), 16–17 November, Athens, Greece.

7. Godwin-Jones, R. (2003) Blogs and Wikis: Environments for Online Collaboration. Language Learning and Technology7(2): 12-16. http://llt.msu.edu/ vol7num2/ pdf/emerging.pdf Last visited July 2011 Stephen Downes, E-Learning 2.0:Notes from Stephen Downeshttp://www.downes.ca/ http://www. speedofcreativity.org/ 2006/10/08/elearning-20-notes-from-stephen-downes/

8. Gradinarova, B. (2009) Importance of web 2.0- mediated competence for the educational demands on learners. In Proceedings of the 10th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing on International Conference on Computer Systems and Technologies (CompSysTech '09),

9. Kirkpatrick, M. (2006) The flu wiki: A serious application of new web tools. http://marshallk.blogspot.com/2005/07/flu-wiki-serious-application-of-new.html (Last visited May 2012)

10. Lamb, B. (2004) Wide open spaces: Wikis, ready or not. EDUCAUSE. Review,39(5) (September/October), 36-48. http://www.educause.edu/pub/ erferm04/erm O452. asp?bhcp=l

11. Leuf, B., Cunningham, W. (2001) The Wiki Way: Quick Collaboration on the Web. Boston, MA: Addison Wesley.

12. Meyer, M. (2010) Creating an Open Web 2.0 Cloud E-Learning Experience, 2010. Educause Annual Conference **http://www.educause.edu/E2010/** Program/SESS124 (Last visited May 2012)

13. Olson, C. (2006) New tools for learning. http://faculty.eicc.edu/golson/ tools.htm (Last visited March 2011)

14. O'Reilly, T. What Is Web 2.0. In Internet: http://www.oreillynet.com/pub/a/ oreilly/tim/news/2005/09/30/what-is-web-20.html

15. Oren Eyal, Max Vëolkel, John G. Breslin& Stefan Decker. (2006) Semantic Wikis for Personal Knowledge Management. S. Bressan, J. KËung, and R. Wagner (Eds.): DEXA 2006, LNCS 4080, pp. 509–518, 2006, Springer-Verlag Berlin Heidelberg http://www.johnbreslin.org/files/publications/ 20060919_dexa2006.pdf
16. Rogers, E. (2011) How Web 2.0 Is Influencing E-Learning (Last visited May 2012)
17. Taylor, C. (2005) It's a Wiki, Wiki World. Time165(23): 40-42.
18. Wheeler, S, Kelly, P. & Gale, K. (2005) The influence of online problem-based learning on teachers' professional practice and identity. ALT-J 2005,3(2):125-137.
19. Wiki, Matrix - Compare Them All. (2007) Retrieved April 19, 2007 from http://www.wikimatrix.org/

Authors:
Teodora Ivanova Bakardjieva, Prof, Ph.D.
Varna Free University, Chaika Resort, Varna 9007, Bulgaria
+ 359 052 359524
bakardjieva@vfu.bg

Antonina Ivanova Ivanova, Assist. Prof.
Varna Free University, Chaika Resort Varna 9007, Bulgaria
+ 359 052 359574
krumova@vfu.bg

COMPUTER SIMULATION AND SYSTEM DYNAMICS
IN EDUCATIONAL PROCESS

Viktors Gopejenko, Yelena Kozmina

Abstract: Specific physical tasks that lead to the systems of two differential equations of the first order are considered in this paper. For each task an imitation model with the help of which numerical solutions of the problem were received is built in Vensim package. Precise solutions of the systems of differential equations are given for the majority of the models. Possibilities of studying the many-sided problem based on Vensim dynamic system modelling package are described.

Key words: System Dynamics, Differential Equation, Computer Simulation, Vensim, Educational Process

Dynamic systems are mathematical models of an object, process or phenomenon, which as opposed to static systems save in the memory their past condition. The past state of the system is connected with the present one by means of a derivative, which is present in the model. The more memory the system has the more past states affect the present and the greater the degree of the higher derivative is used to describe the model.

Without relevant experience, theory is uninteresting to students. When experiments in the real world are impossible, simulation becomes the main way we can learn effectively about the dynamics of complex systems [4]. The modeling exercises guide the learner through the process of building a working simulation; students will not only learn about the issues addressed, and in the use of state of the art simulation software, but will develop skill in the modeling process.

Vensim is a package that provides dynamic modeling [1, 2, 3]. The main advantages of the Vensim package are the possibilities to build cause-effect diagrams, flow diagrams, tree-like relationships, to create documentation for the model, to identify the relationship cycle, to use the expression editor, to use built-in functions, to check units of measurement, to trace conditions, to check a chart, to carry out an experiment, to plot graphs that allow to determine the sensitivity to the variable parameters, to generate tabular reports, to compare the results of experiments.

Let's consider the mathematical models, quantitative and qualitative characteristics of the behavior of bodies moving near the Earth's surface.

The fall of a body is the movement of a body in the Earth's gravitational field with initial velocity equal to zero. We perceive the body as a material point. If you direct a coordinate Oy axis down vertically

aligning the point of origin to the original position of the body (Fig. 1), then the coordinate $y(t)$ is the path traversed by a point from the initial position.
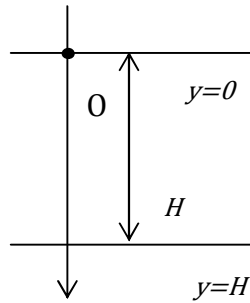

Fig. 1. The axis of coordinates for the body falling from the height H

Using the physical meaning of the derivative we define the relation between $y(t)$, velocity $v(t)$ and acceleration $a(t)$

$$\begin{cases} dv(t)/dt = a(t), \\ dy(t)/dt = v(t). \end{cases} \quad (1)$$

According to the second Newton law: $a(t) = F/m$, where m is the weight of the body, $F = F(t, v, y)$ is the force acting on the body, $a(t)$ is acceleration transmitted by that force. The solution of (1) has the form:

$$\begin{cases} v(t) = v_0 + a \cdot (t - t_0), \\ y(t) = y_0 + v_0 \cdot (t - t_0) + a \cdot (t - t_0)^2/2, \end{cases} \quad (2)$$

where y_0 and v_0 denote a coordinate and velocity of a material point at time t_0. In particular, if $t_0 = 0, v_0 = 0, y_0 = 0$, then

$$\begin{cases} v(t) = a \cdot t, \\ y(t) = a \cdot t^2/2. \end{cases} \quad (3)$$

Model 1. The falling of the body without the resistance of the environment.

Consider the free fall of the body, i.e. the falling of the body on the ground under gravity without air resistance; g is the acceleration of the free fall. In the case of the free fall of the body the solution of the system (1) is as follows:

$$\begin{cases} v(t) = g \cdot t, \\ y(t) = g \cdot t^2/2. \end{cases} \quad (4)$$

306

If a body falls from a height H, then $y(t) = H$ with $t = \sqrt{2H/g}$. The body at the moment of landing has the velocity $v = \sqrt{2gH}$.

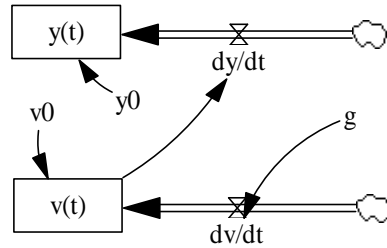Let's construct a model of *Vensim PLE* for the system of differential equations (1), suggesting that a(t)=g.


```
"dv/dt"=g
"dy/dt"="v(t)"
FINAL TIME = 4
g=9.807
INITIAL TIME = 0
TIME STEP = 0.125
"v(t)"= INTEG ("dv/dt",v0)
v0=10
y(t)"= INTEG ("dy/dt",y0)
y0=0
```

Fig. 1.a A model of *Vensim PLE* for the system (1)

Fig. 1.b. Equations of the model with

$$y_0 = 0, v_0 = 0, g = 9.807.$$

Model 2. The fall of the body with the resistance of the environment.

In the problem of falling bodies let's take into account the decelerating force due to air resistance. The fall of the body under the gravity force $\overrightarrow{F_g}$ interferes with the force of the resistance of the environment $\overrightarrow{F_d}$ which is dependent of the velocity of movement of the body (Fig. 3). The value of the resulting force acting on a material point is calculated as follows:

$$F = F_g - F_d = mg - F_d, \tag{5}$$

where $F_g = mg$ is the value of the gravity force, $F_d = F_d(v)$ is the value of resistance force of the environment.
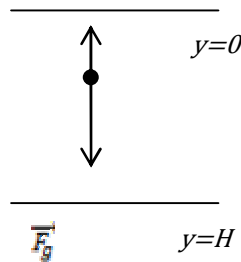

Fig. 3. Vectors of the forces that act on the falling body (taking into account air resistance)

Assumed dependency F_d on the velocity $v(t)$ is considered to be experimentally confirmed, if the values found for $y(t)$ are consistent with the experimental ones.

The force of the resistance is proportional to the velocity of the moving body.

Suppose that $F_d = k_1 \cdot v(t)$, where the parameter k_1 depends on the geometry of the body and the properties of the environment. This dependency is not a law of physics, however, it is a phenomenological expression which is valid in a limited range of velocities. Thus, the system (1) can be written as

$$\begin{cases} dv(t) = g(1 - k_1 \cdot v(t)/(mg))dt, \\ dy(t) = v(t)dt. \end{cases} \tag{6}$$

Since the function $F_d(v)$ is increasing, there is a limit or the steady velocity of $v_1 = mg/k_1$ that corresponds to the condition $mg = F_d$ and zero acceleration.

In the case of the free fall with air resistance which is directly proportional to the velocity of the fall, the solution of the system (1) has the form:

$$\begin{cases} v(t) = v_1\left(1 - e^{-gt/v_1}\right), \\ y(t) = v_1 t + v_1^2 \cdot \left(e^{-gt/v_1} - 1\right)/g. \end{cases} \tag{7}$$

Remarks. When $k_1 \to 0$, the solutions $v(t)$ and $y(t)$ defined by (7) tend to solutions of (4) that are obtained in the case of the body falling without the resistance of the environment.

Let's construct a *Vensim PLE* model for the system of differential equations (6) assuming that $m = 10^{-2}\,kg$ and $k = 10^{-4}\,kg/m$.
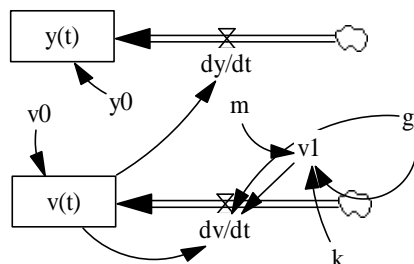

```
"dv/dt"=g*(1-"v(t)"/v1)
"dy/dt"="v(t)"
FINAL TIME = 16
g=9.807
INITIAL TIME = 0
k=0.0001
m=0.01
SAVEPER = TIME STEP
TIME STEP = 0.125
"v(t)"= INTEG ("dv/dt",v0)
v0=0
v1=m*g/k
"y(t)"= INTEG ("dy/dt",y0)
y0=0
```

Fig. 2.a A model of *Vensim PLE* for the system (6).

Fig. 2.b. Equations of the model

308

Next, we compare the results obtained by taking into account the air resistance proportional to the velocity of the moving body, and the results obtained without the resistance.
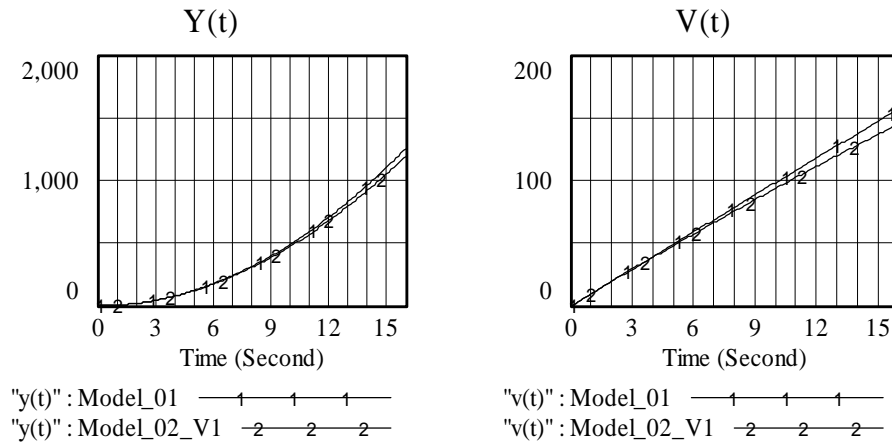

Fig. 5. Dependences $y(t)$ and $v(t)$ of the free fall of the body without air resistance (Model 01) and with air resistance (Model 02 V1) proportional to the velocity of the moving body.

The force of the resistance is proportional to the square of the velocity of the moving body.

We assume that dependence of the resistance force on velocity is quadratic

$F_d(v) = k_2 \cdot v^2$, where k_2, is the coefficient of proportionality; generally speaking, the coefficient k_1 is different in case when resistance of the environment is proportional to the velocity. The top or steady velocity v_2 corresponding to the condition $mg = F_d$ and zero acceleration may be calculated as follows: $v_2 = \sqrt{mg/k_2}$. Velocity $v(t)$ may be measured in the same units as the top velocity. Let's transform F_d to: $F_d = mg(v/v_2)^2$.

The force acting on a falling body has the form: $F(v) = mg(1 - (v/v_2)^2)$, and the system (1) is written down as:

$$\begin{cases} dv(t)/dt = g(1 - (v/v_2)^2), \\ \qquad dy(t)/dt = v(t). \end{cases} \qquad (8)$$

From the first equation of the system (8) we can get the following: $v(t) = v_2\left(1 - 2/\left(e^{2gt/v_2} + 1\right)\right)$. Let's place $v(t)$ into the second

equation of the system (8), then, integrate, and find $y(t)$: $y(t) = y_0 + v_2{}^2 \cdot \left(\ln\left(e^{2gt/v_2} + 1 \right) - gt/v_2 - \ln 2 \right)/g$.

Assuming that $y_0 = 0$ in case of the free fall of the body with the resistance force, which is proportional to the square of the velocity of the moving body, the solution of the system (8) has the following form:

$$\begin{cases} v(t) = v_2\left(1 - 2/\left(e^{2gt/v_2} + 1\right)\right), \\ y(t) = v_2{}^2 \cdot \left(\ln\left(e^{2gt/v_2} + 1\right) - gt/v_2 - \ln 2 \right)/g. \end{cases} \tag{9}$$

Remarks. When $k_2 \to 0$, the solutions $v(t)$ and $y(t)$ defined by (9) tend to solutions (4) that are obtained in case of the body falling without the resistance of the environment.

Let's construct a *Vensim PLE* model for the system of differential equations (8) assuming that $m = 10^{-2}\, kg$ and $k_2 = 10^{-4}\, kg/m$.

```
"dv/dt"=g*(1-"v(t)"*"v(t)"/(v2*v2))
"dy/dt"="v(t)"
FINAL TIME = 4
g=9.807
INITIAL TIME = 0
k2=0.0001
m=0.01
TIME STEP = 0.125
"v(t)"= INTEG ("dv/dt",v0)
v0=0
v2=SQRT(m*g/k2)
"y(t)"= INTEG ("dy/dt",y0)
y0=0
```
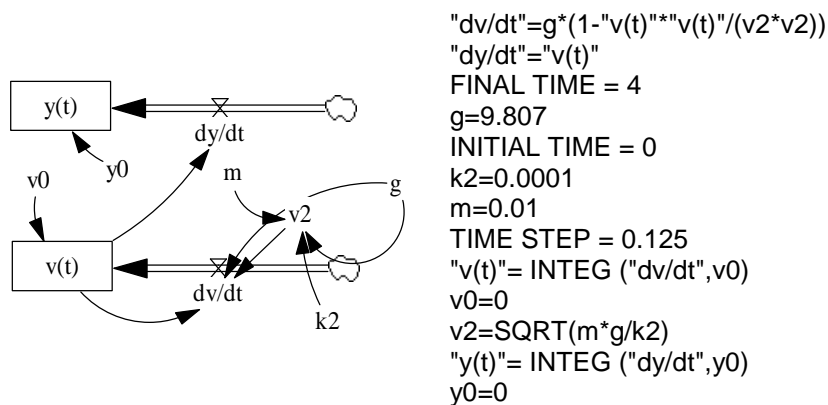
Fig. 6.a A model of *Vensim PLE*   Fig. 2.b. Equations of the model
for the system (8)

Let's compare the results received taking into account air resistance which is proportional to the square of the velocity of the moving body with the results received without air resistance (Fig. 7). Note that the plots of the $y(t)$ and $v(t)$ functions received in the *Vensim PLE* model in the case of the free falling body taking into account air resistance match the corresponding plots of the exact solutions (9) of the system (8).

Fig. 7. Dependencies $y(t)$ and $v(t)$ of the body free fall without inclusion of the air resistance (Model 01) and taking in to account air resistance (Model 02 V2) which is proportional to the square of the velocity of the moving body.

Model 3. The movement of the body thrown up vertically (without air resistance).

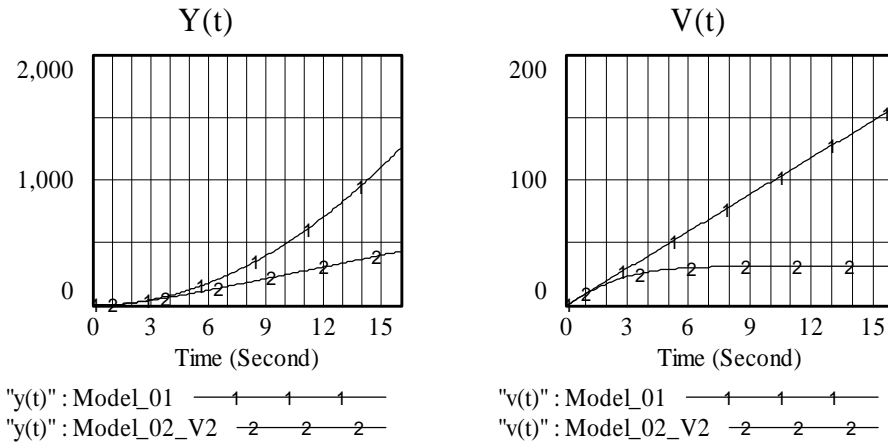Differential equation system (1) and its solution (2) may be used in the task of the body movement that is thrown up vertically with a certain initial velocity v_0 without air resistance. In this case it is convenient to direct Oy axis up vertically and to match its beginning with the point of throwing. In the equations of the uniform accelerated in-line motion it is necessary to consider $y_0 = 0, a = -g, v_0 > 0$. Equation (10) follows

from (2):

$$\begin{cases} v(t) = v_0 - g \cdot t, \\ y(t) = v_0 \cdot t - g \cdot t^2/2. \end{cases} \tag{10}$$

In time $t = v_0/g$ the velocity of the body $v(t)$ becomes zero, i.e. the body achieves the maximum point of the ascent. The body returns to the ground ($y(t) = 0$) in time $t = 2v_0/g$, consequently the times of the ascent and of the descent are equal. When body falls on the ground its velocity is equal to $v(2v_0/g) = -v_0$, i.e. the body falls on the ground with the same absolute velocity that it was thrown up. The maximum height of the ascent is $y(v_0/g) = v_0{}^2/(2g)$. In the *Vensim PLE* model 1 for the system (2) let us assume that $y_0 = 0, v_0 = 60, a = -g$. Let us draw up the dependencies $y(t)$ and $v(t)$.
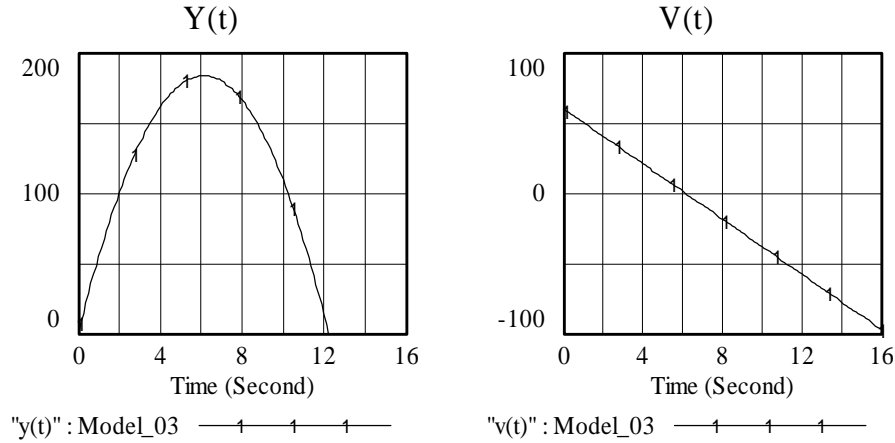
Y(t)

V(t)

"y(t)" : Model_03 ——┼———┼———┼——   "v(t)" : Model_03 ——┼———┼———┼——

Fig. 8. Dependencies $y(t)$ and $v(t)$ of the body movement up and down without air resistance $y_0 = 0$, $v_0 = 60$, $g = 9.807$.

Model 4. The movement of the body thrown up (taking into account air resistance).

The gravitation force directed downwards and resistance force directed in direction opposite to the movement, i.e. also downwards affect the body. The velocity vector is directed upwards, the movement is decelerated, and consequently the acceleration vector is also directed downwards. Then $dv(t)/dt = -g - F_d/m.$ Let's consider that decelerating force F_d is proportional to the square of the velocity: $F_d = k \cdot v^2$. Let's introduce coefficient $k_2^2 = k/(mg)$, then $F_d = mg \cdot k_2^2 \cdot v^2$. Then system of equations (11) is as follows:

$$\begin{cases} dv(t)/dt = -g\left(1 + k_2^2 \cdot v^2(t)\right), \\ dy(t)/dt = v(t). \end{cases} \tag{11}$$

with the initial condition $v(0) = v_0, y(0) = 0.$

When the body starts falling the velocity vector is directed downwards the movement is accelerated, consequently the acceleration vector is also directed downwards. The gravitation force is directed downwards, the resistance force – upwards, consequently:

$$\begin{cases} dv(t)/dt = -g\left(1 - k_2^2 \cdot v^2(t)\right), \\ dy(t)/dt = v(t). \end{cases} \tag{12}$$

312

Integrating the first equation of the system (11), substituting the received $v(t)$ in the second equation of the system (11), integrating in and receiving the solution of the equation (11):

$$\begin{cases} v(t) = (1/k_2) \cdot (k_2 \cdot v_0 - tg(gk_2\, t))/(1 + k_2 \cdot v_0 \cdot tg(gk_2\, t)), \\ \qquad y(t) = ln|cos(gk_2\, t) + k_2 \cdot v_0 \cdot sin(gk_2\, t)|/(gk_2^2). \end{cases} \tag{13}$$

At $t_1 = \operatorname{arctg}(k_2 \cdot v_0)/(k_2 \cdot g)$ the velocity $v(t_1) = 0$. From this moment the body starts falling, the movement of the body is described by the system of equations (12) with the initial conditions $v(t_1) = 0, y(t_1) = ln(1 + k_2^2 v_0{}^2)/(2gk_2^2)$. The solution of this system is the following:

$$\begin{cases} \qquad v(t) = (1/k_2) \cdot \left(1 - e^{2k_2 \cdot g \cdot (t - t_1)}\right)/\left(1 + e^{2k_2 \cdot g \cdot (t - t_1)}\right), \\ y(t) = \left(ln\left(2\sqrt{1 + k_2^2 v_0^2}\right) + gk_2(t - t_1) - ln\left(1 + e^{2k_2 \cdot g \cdot (t - t_1)}\right)\right)/(gk_2^2). \end{cases} \tag{14}$$

Let's build *Vensim PLE* model for the systems (11) and (12) assuming that $m = 10^{-2}\, kg$, $k_2 = 10^{-4}\, kg/m$.


```
"dv/dt"=-g+IF THEN ELSE(
"v(t)">=0,
-g*"v(t)"*"v(t)"*k2*k2,
g*"v(t)"*"v(t)"*k2*k2)
"dy/dt"="v(t)"
FINAL TIME = 16
g=9.807
INITIAL TIME = 0
k=0.0001
k2=SQRT(k/(m*g)
m=0.01
TIME STEP = 0.125
"v(t)"= INTEG ("dv/dt",v0)
v0=60
v2=SQRT(m*g/k2)
"y(t)"= INTEG (IF THEN
ELSE("y(t)">=0, "dy/dt", 0), y0)
y0=0
```
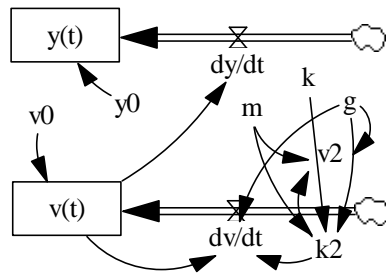
Fig. 9.a. *Vensim PLE* model of the movement of the body thrown up vertically taking with air resistance proportional to the square of the velocity

Fig. 9.b. Equations of the model

313

Note that transformation from the system (11) to system (12) in the equations of the model occurs with the logical operator that significantly simplifies the solution of the task.

The plots of the functions $y(t)$ and $v(t)$ are the following:
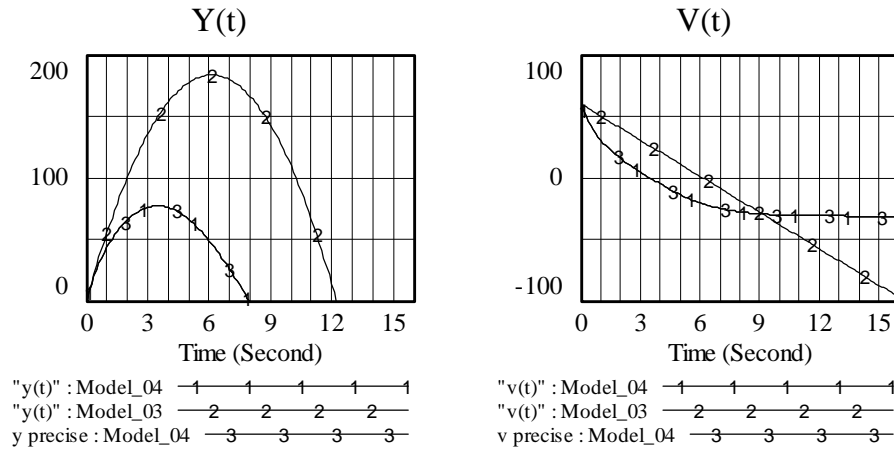
Fig. 10. Dependencies $y(t)$ and $v(t)$, of the body movement up and down $y_0 = 0$, $v_0 = 60$, $g = 9.807$. Plots for $y\ precise$ and $v\ precise$ correspond to the precise solutions of the equation systems (13) and (14)

Model 5. The movement of the body thrown at an angle to the horizon (without air resistance).

The body is thrown at an angle to the horizon with the initial velocity $\overrightarrow{v_0}$. Let's define the trajectory of the body movement without air resistance. Let's consider that the movement of the body occurs in one plane. Let's place the origin of the coordinates O in the point of the start of the body movement. The velocity vector $\vec{v} = \vec{v}(t)$, directed tangentially to the body trajectory consists of two components $v_x = v_x(t)$ and $v_y = v_y(t)$. The force $\vec{F}$ affecting the body is also decomposed on two components (F_x, F_y). In this case the movement in the plane is the result of two movements: along Ox axis with the velocity v_x under the force F_x and along Oy axis with the velocity v_y under the force F_y (Fig.11)

314
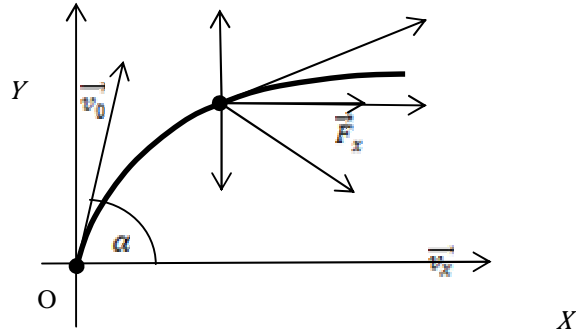
Fig. 11. Vectors of the forces and of the velocity that act on the body (without air resistance)

According to the second Newton law the force and acceleration (derivative from the velocity) are related as follows:

$$\begin{cases} m \cdot dv_x/dt = F_x, \\ m \cdot dv_y/dt = F_y. \end{cases} \tag{15}$$

As the air resistance is neglected, considering that the body is affected only by the gravitation force directed towards the ground, than $F_x = 0, F_y = -mg.$ In this case the system of the equation is as follows:

$$\begin{cases} dv_x/dt = 0, \\ dv_y/dt = -g, \end{cases} \tag{16}$$

with the following solution

$$\begin{cases} v_x = C_1, \\ v_y = -gt + C_2, \end{cases} \tag{17}$$

where C_1, C_2 – arbitrary constants.

Let's assume α – is an angle that forms vector $\overrightarrow{v_0}$ with positive direction of the Ox axis, $v_0 = |\overrightarrow{v_0}|$, then initial conditions

$$\begin{cases} v_x(0) = v_0 \cdot \cos \alpha, \\ v_y(0) = v_0 \cdot \sin \alpha \end{cases} \tag{18}$$

allow defining C_1, and C_2:

$$\begin{cases} C_1 = v_0 \cdot \cos \alpha, \\ C_2 = v_0 \cdot \sin \alpha, \end{cases} \tag{19}$$

receiving

$$\begin{cases} v_x = v_0 \cdot \cos \alpha, \\ v_y = -gt + v_0 \cdot \sin \alpha. \end{cases} \tag{20}$$

As $v_x = dx/dt$ and $v_y = dy/dt$, where $x = x(t)$ is the displacement of the body along Ox axis, $y = y(t)$ is the displacement of the body along Oy axis, then

$$\begin{cases} dx/dt = v_0 \cdot \cos\alpha, \\ dy/dt = -gt + v_0 \cdot \sin\alpha. \end{cases} \tag{21}$$

The solution of this system looks as follows

$$\begin{cases} x = v_0 \cdot \cos\alpha \cdot t + C_3, \\ y = v_0 \cdot \sin\alpha \cdot t - gt^2/2 + C_4. \end{cases} \tag{22}$$

Considering that $x(0) = 0, y(0) = 0$, $C_3 = C_4 = 0$ and, consequently,

$$\begin{cases} x = v_0 \cdot \cos\alpha \cdot t, \\ y = v_0 \cdot \sin\alpha \cdot t - gt^2/2. \end{cases} \tag{23}$$

Expressing t from the first equation of the system $t = x/(v_0 \cdot \cos\alpha)$ and substituting in the second equation, function $y = x \cdot \text{tg}\,\alpha - g \cdot x^2/(2v_0^2 \cdot \cos^2\alpha)$ is received, the plot of which is parabola.

Remark: Vensim possibilities allow changing the initial conditions and the parameters of the problem observing the changes in solutions, i.e. to dynamically visualize the latter.
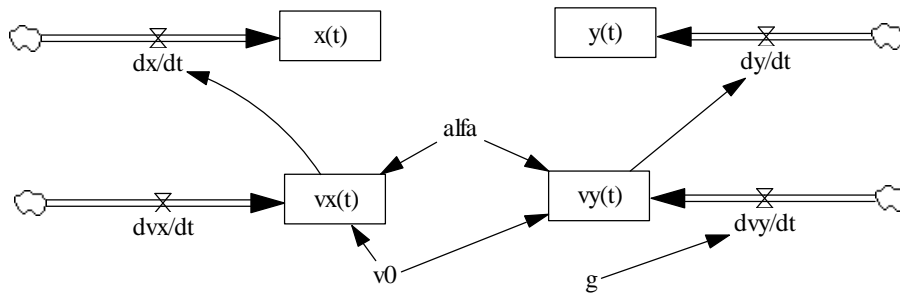
Fig. 12.a A model of *Vensim PLE* for the system (16), (21)
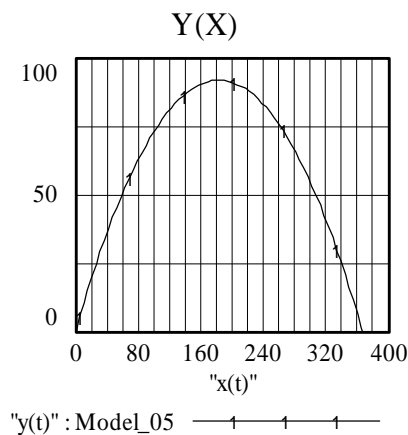
Fig.12.b Movement of the body
thrown at an angle to the horizon

$y_0 = 0, \; v_0 = 60, \; alfa = 45$.

alfa=45
"dvx/dt"=0
"dvy/dt"=-g
"dx/dt"="vx(t)"
"dy/dt"="vy(t)"
FINAL TIME = 16
g=9.807
INITIAL TIME = 0
SAVEPER = TIME STEP
TIME STEP = 0.125 v0=60
"vx(t)"= INTEG
("dvx/dt",v0*cos(alfa*3.1415/180))
"vy(t)"= INTEG
("dvy/dt",v0*sin(alfa*3.1415/180))
"x(t)"= INTEG ("dx/dt",0)
"y(t)"= INTEG ("dy/dt",0)

Fig.12.c Equations of the model

Model 6. Movement of the body thrown at an angle to the horizon (taking into account air resistance)

Let's consider the relations between the coordinates $v_x(t)$ and $v_y(t)$ of the velocity vector $\vec{v}$ in the case when the body is thrown at an angle α to the horizon taking into account air resistance. In each moment of time the body position is described by the coordinates (x, y), where $x = x(t), y = y(t)$. Velocity vector $\vec{v} = \vec{v}(t)$ with absolute value of $v(t) = \sqrt{v_x^2(t) + v_y^2(t)}$ is directed tangentially to the body movement trajectory and at the initial time moment is equal to $\vec{v_0}$

Let's assume $\theta = \theta(x, y)$ – is the angle that forms vector $\vec{v}(t)$ with positive direction of the Ox axis, $\vec{F_d} = (-F_d \cdot cos\theta, -F_d \cdot sin\theta)$, – deceleration force directed opposite to the velocity vector $\vec{v}$, $F_d = |\vec{F_d}|$ (Fig.13)

317
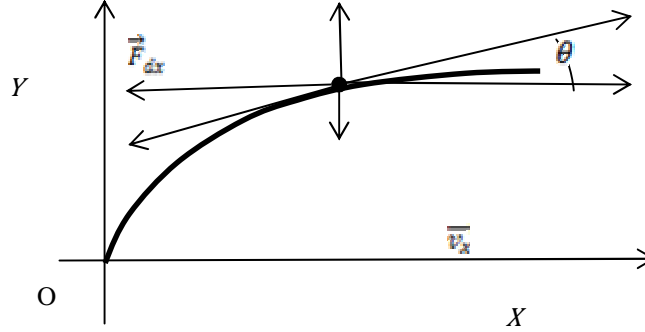
Fig. 13. Vectors of the forces and of the velocity that act on the body (with air resistance).

Then according to the second Newton law,

$$\begin{cases} m \cdot dv_x(t)/dt = -F_d \cdot \cos\theta, \\ m \cdot dv_y(t)/dt = -m \cdot g - F_d \cdot \sin\theta. \end{cases} \tag{24}$$

As

$$\begin{cases} v_x(t) = v(t) \cdot \cos\theta, \\ v_y(t) = v(t) \cdot \sin\theta, \end{cases} \tag{25}$$

then

$$\begin{cases} m \cdot dv_x(t)/dt = -F_d \cdot v_x(t)/v(t), \\ m \cdot dv_y(t)/dt = -m \cdot g - F_d \cdot v_y(t)/v(t). \end{cases} \tag{26}$$

Assuming that the absolute value of the decelerating force is proportional to the square of the absolute value of the velocity, i.e.

$$F_d = k_2 \cdot v^2(t), \tag{27}$$

receiving $F_d/v(t) = k_2 \cdot v(t) = k_2 \cdot \sqrt{v_x^2(t) + v_y^2(t)}.$

As a result of these transformation equation system (24) is written as follows

$$\begin{cases} m \cdot dv_x(t)/dt = -k_2 \cdot \sqrt{v_x^2(t) + v_y^2(t)} \cdot v_x(t), \\ m \cdot dv_y(t)/dt = -m \cdot g - k_2 \cdot \sqrt{v_x^2(t) + v_y^2(t)} \cdot v_y(t). \end{cases} \tag{28}$$

Note that the right sides of the equations contain the absolute value of the velocity $\sqrt{v_x^2(t) + v_y^2(t)}$, therefore it is impossible to find one of the components of the velocity vector not knowing the other one. System of equations

$$\begin{cases} dx(t)/dt = v_x, \\ dy(t)/dt = v_y \end{cases} \tag{29}$$

318

links the coordinates $x(t), y(t)$ of the moving point with the components of the velocity vector $v_x(t)$ and $v_y(t)$.

Let's construct the *Vensim PLE* model for the system of the differential equations (24), (29) with the decelerating force (27).
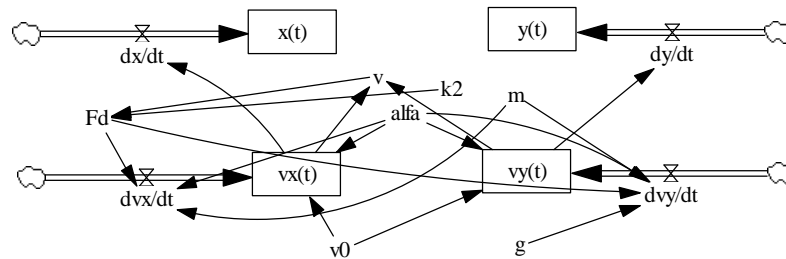
Fig. 14.a. A model of *Vensim PLE* for the systems (24), (29*)*


```
alfa=45
("dvx/dt"=-
Fd*cos(alfa*3.1415/180)/m
"dvy/dt"=-g-
Fd/m*sin(alfa*3.1415/180)
"dx/dt"="vx(t)"
"dy/dt"="vy(t)"
Fd=k2*v*v
FINAL TIME = 16
g=9.807
INITIAL TIME = 0
k2=0.0001
m=0.01
SAVEPER = TIME STEP
TIME STEP = 0.125
v=SQRT("vx(t)"*"vx(t)"+"vy(t)"*"vy(t
)")
v0=60
"vx(t)"= INTEG
("dvx/dt",v0*cos(alfa*3.1415/180))
"vy(t)"= INTEG
("dvy/dt",v0*sin(alfa*3.1415/180))
"x(t)"= INTEG ("dx/dt",0)
"y(t)"= INTEG ("dy/dt",0)
```
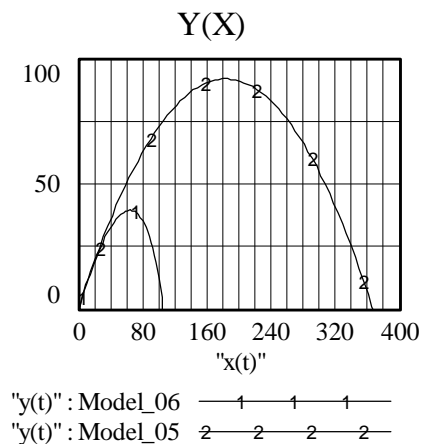
Fig.14.b. The movement of the body thrown at an angle to the horizon $y_0 = 0, \; v_0 = 60, \; alfa = 45, m = 0.$

Fig.14.c. Equations of the model

Conclusion

The result of simulated modeling in Vensim package is a solution that is close enough to the exact solution of differential equations. It means that such approach makes it possible to identify the most stable relationships that directly affect the solution of the problem. The main advantage of the Vensim package is the possibility to visualize the solutions of the differential equations dynamically.

References
1. Harley Gould, Jan Tobochnik. (1988) An Introduction to Computer Simulation Methods Applications to Physical Systems. Addison-West Publishing Company. pp. 349.
2. Juan Martin Garcia. (2006) Theory and Practical Exercises of System Dynamics. Barcelona: Juan Martin Garcia, p.294.
3. V. Gopejenko, Y. Kozmina. (2012) System Dynamics and System Thinking in Educational Process. The 10th International Conference Information Technologies and Management 2012. April 12-13, 2012, Information Systems Management Institute, Riga, Latvia, p.51-52
4. Vensim Ventana. Simulation Environment (2007), DSS, Professional, PLE Plus, PLE. With Causal Tracing Reality Check and SyntheSim User's Guide Version 5. 1998-2007 Ventana Systems, Inc. Revision Date: July 4, 2007, p.299.

Authors:
Dr. Ing. **Viktors Gopejenko**, Dr. Math. **Yelena Kozmina**
Information Systems Management Institute
Department of Natural Sciences and Computer Technologies
Ludzas 91, LV-1019, Riga, Latvia,
e-mail: viktors.gopejenko@isma.lv; jelena.kozmina@isma.lv

CREATING NEW AVENUES OF EDUCATION USING A VIRTUAL WORLD SECOND LIFE

Yanka Shurelova

Abstract: *Virtual worlds provide learners with a unique learning environment and teaching. Learners can interact with environments and personalities in ways that can be difficult to manage for a large group in traditional training. With the growing popularity of multi-user virtual environments such as Second Life, learners can now be engaged in ways that are not possible outside the virtual world. With the ability to defy the laws of physics and "real-world" realities and as the technology behind Second Life evolves and matures, the face of teaching, learning and research is evolving and adapting. We particularly take into consideration the relevant aspects of teaching and popularization of science in 3D virtual environments.Virtual environments such as Second Life seem to be an obvious vehicle for the transfer of knowledge that has the potential to appeal to the millennial learning style. The virtual world Second Life enhances the educational success of 21st century learners.*

Keywords: Virtual worlds,Second life,Virtual learning, Science, Teaching, Stars, Hertzshprung – Russell Diagram

We live in a world of risk and interaction of complex systems. A world in which we talk about global warming, as well as "freezing" of the global economy. Natural systems are now dependent on human beings. Things that were once "creation of the Creator" are now human affairs. The new millennium created a new phenomenon in society - social networks, blogs and so-called "virtual worlds". Since the advent of virtual worlds has become extremely popular it quickly won a large base of regular users. The areas where this technology successfully and actively apply: design and manufacturing, military, medicine, architecture and in recent years and education. Applications of virtual worlds opens enormous potential for consumers and students, lecturers and teachers, politicians and managers to implement cooperation in 2 D or 3 D environment. In my opinion, in the twenty-first century requires new knowledge and skills that we will not get them in schools alone and we will never get them in the schools we currently have.

What is a virtual world or 3D virtual environment? The term "virtual world" is used by many researchers and writers working in the field of development of simulations and virtual realities. The lack of common definition is the reason the term be used in different ways in different contexts by academics, actors and professionals in business and media representatives [6, 7, 10, 17]. A review of many existing definitions allows us to find common ground between them and it is: "The virtual world is an

online community - computer-based simulated environment in which users can interact with each other to use and create objects. Virtual worlds - interactive 3D virtual environment allows users to experience through their avatars, text, 2D and 3D graphics performance and communicate through text, graphical icons, visual gesture and sound. "[4]

Contemporary virtual world technology is comprised of more than the computer software and hardware that enables numerous users to simultaneously share, compare, and construct new knowledge, information and experiences. Whether for the purposes of playing games, constructing simulations, or expanding the possibilities of how, what, who, and when we teach, contemporary virtual world technology also brings new ways to consider the parallels between social networks and social learning environments. Further, online virtual worlds may be the most expansive form of social communication media in which users share information and simultaneously consider its limits. In the context of education this means that educators and learners are constantly shaping and redefining not simply what, when, where, and how education could be but rather how it is now.

Why are we using virtual worlds? Kids and games are two things that constantly go together. Computer games are some of the most convenient and funny ways for kids to play games. Virtual World games can reach a wide range of students on a wide range of subjects. By interacting and communicating with other users, students can practice real world skills in a virtual place. Whether you have a few minutes left in the classroom or you want to give student free time in the computer lab, these games are sure to keep your students' attention! And begin training in virtual environments – like a game.

Who is using virtual worlds? Although virtual worlds have been around for over 20 years, their real potential and use for educational purposes has grown in the last 5 years, including replicating universities, museums, art galleries and science labs for tutoring and mentoring[4,24]. According to the article Serious Virtual Worlds [14] "…the lines between virtual worlds, games and social networking are blurring significantly leading to the assertion that over the next 5 years the majority of young people under 18 will have avatars and be using these kinds of applications daily and therefore have different expectations about how education may be delivered to them".

Which virtual worlds? Lack of unique classification explains the creation of a variety of topologies for the needs of specific targets. According to Dr. Maria Monova – Jeleva in terms of opportunities and potential for implementation of training different types of software applications for generating virtual worlds are grouped into five categories:

• Game Worlds (Worlds of Role Play Games) such as World of Warcraft, Everquest, Guild Wars;

• Social worlds such as Second Life, Cy World, Active Worlds, Hobbo Hotel, club Penguin, Pandanda;

• Corporate / business worlds such as Wonderland, Metaverse of IBM;

• Worlds for learning such as America's Army, platform OLIVE;

• Mirror worlds such as Google Earth, Virtual Earth, NASA World Wind.[1]

The following[Table 1.] is a comparison table of virtual worlds of VEIG - Virtual Environment Info Group - Association of experts, professionals and developers in virtual worlds [36].

Table 1. Comparison table of virtual worlds

Comparison of Virtual Environments									
Virtual World	OS	Cost per month	Target user & style	Edit avatars?	Build or design content?	Script content?	Own land or sell items?	Education ready?	Comm. Events?
Active Worlds	PC & Linux	Free / $6.95	General; Exploration	✓	✓	✓	✓ No selling	Can code	✓
Barbie Girls	PC	Free	Young girls; Fashion, social	✓	✓	No	Neither	No	✓
Club Penguin	PC & Mac	Free / $5.95	Kids; Games and Activities	✓	No	No	Neither	No	✓
Forterra Systems	PC	Contract	Training, E-Learning, Serious Games	✓	Can code	✓	✓	Can code	Can code
Gaia Online	PC & Mac	Free	Social; Top-down overview, sprites	✓	No	No	Sorta	No	✓
Habbo Hotel	PC & Mac	Free	Teens; Social	✓	✓	No	Neither	No	✓
Kaneva	PC & Mac	Free	Teens; Social	✓	✓	No	Sorta	No	✓
Neopets	PC & Mac	Free / $7.95	Kids and teens; Mini-games, social	✓	No	No	✓✓	No	✓
Teen Second Life	PC & Mac	Free / $9.95	Teens only; 3D, Creation, social	✓	✓	✓	✓✓	Can code	✓

Second Life	PC & Mac	Free / $9.95	18+ only; 3D, Creation	✓	✓	✓	✓	Can code	✓
The Sims Online	PC	$9.95/mo.	General; Social; 3D	✓	✓	No	✓ No selling	No	✓
There	PC	Free / $9.95	General; Social	✓	✓	Limited	✓	No	✓
Webkins	PC	Free	Social	✓	No	No	Neither	No	✓
Whyville	PC & Mac	Free	Kids and teens; 2D sprites; Educational	✓	✓	No	Neither	✓	✓
Zwinktopia	PC & Mac	Free	Teens; Social	✓	No	No	Neither	No	✓

Presently most of the expertise of the developing team is focused on Second Life and its derivates.   The only available open source projects that might eventually be used for this kind of educational application are presently based (more or less loosely) on Linden Labs technology. Most of the research that has been done so far in this field has used Second Life or a derivate as its test bed.

What is Second life? Second Life ®[Figure 1.][28] is "a 3D virtual world created by its Residents (people like you) that's bursting with entertainment, experiences, and opportunity. Second Life offers a uniquely immersive experience where you can create, buy, and sell anything you can imagine; socialize with people across the world; and enjoy live events and gaming activities. Headquartered in San Francisco, we're focused on making the virtual world a better place for our Second Life community. If you're smart, talented, and looking for ways to make a difference in the virtual world, we invite you to explore our exciting employment opportunities". Founded in 1999, Linden Lab is the maker of Second Life [19].

Figure 1. www.secondlife.com

Second Life is intended for people aged 16 and over, and from 2011 there are about one million active users with avatars[Figure 2.]. [25]. Built into the software is a three-dimensional modeling tool based on simple geometric shapes that allows residents to build virtual objects [Figure 3.1.]. There is also a procedural scripting language, Linden Scripting Language [Figure 3.2.], which can be used to add interactivity to objects. Sculpted prims (sculpties), mesh, textures for clothing or other objects, and animations and gestures can be created using external software and imported.
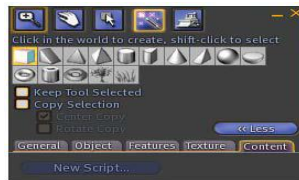
Figure 2. Avatar   Figure 3.1. Building tools   Figure 3.2. Linden Scripting Language
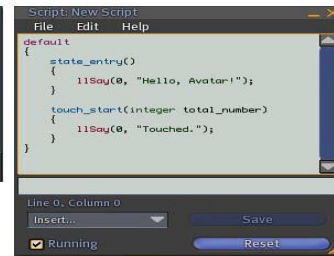
In Second Life, one of the aims of the virtual world is to socialize, making this quite unusual in gaming strategies. More specifically, Second Life is a multi-user virtual environment or a virtual world with a rich social network. It has evolved from the context of computer based games and is part of the current explosion in the use of Web 2.0 technologies. It is a commercial enterprise created by Linden Labs in 2003 selling and renting virtual land, as well as managing a virtual economy and providing a level of administration, policing and censorship [20]. A portal is also provided for educators to exchange ideas and for researchers to network. This virtual world is free to use, although many users pay money (called Lindens in the virtual world) to purchase items virtually. Linden Labs also provides a secure environment for under age users called Teen Second Life. One of the most recent educational studies using Teen Second Life is the Schome pilot project report from the Open University in the United Kingdom. This project in Teen Second Life investigated the level of engagement of 149 high school students [31].

Gartner, Inc. [15] estimates that by 2011, 80 percent of active internet users, including Fortune 500 enterprises, will have a "second life"

in some form of virtual world environment. It also seems clear, however, that virtual worlds in whatever form will be a widely used knowledge and social-interaction tools and will become another part of the social-technical system people use for teaching and learning in the foreseeable future[21].

Researchers report on the limitations of using virtual worlds, particularly technical limitations, including latency. Brainbridge [5] suggests that, in general, the deficiencies of the Internet are highlighted when using Second Life. This is, in part, because of bandwidth problems causing latency or lag in downloading times. This is particularly noticeable in countries such as Australia that has slower internet speeds than other developed countries. A recent paper by Fetscherin and Latteman [13] found that even with improving technology available there are still technical challenges for users of Second Life. Despite the problems in using Second Life that have been reported, many feel that these problems can be overcome and the benefits make it well worth the effort. One of the positives of Second Life is that it allows the instructor or teacher to create educational simulations and games [22]. Moreover, one advantage that Second Life has over other virtual worlds is that all users can create content, as it is not limited to the teacher.

Why education in Second life? Millennials, the generation born after 1980, are the first generation that grew up surrounded by digital media during their formative years of childhood. Virtual environments such as Second Life seem to be an obvious vehicle for the transfer of knowledge that has the potential to appeal to the millennial learning style. The virtual world of Second Life is a global community of creativity, collaboration, commerce and entertainment in which one creates an "avatar" that interacts in a virtual 3D environment; with the million of resident accounts . The Second Life browser software comes equipped with build-tools, group options, communication tools such as chat and instant messenger and tools for users to record "machinima" of their experiences, simulations or live events. With the growing popularity of multi-user virtual environments such as Second Life, learners can now be engaged in ways that are not possible outside the virtual world. With the ability to defy the laws of physics and "real-world" realities, and as the technology behind Second Life evolves and matures, the face of teaching, learning, and research is evolving and adapting.

Teachers and university administrators are experimenting with a new form of virtual learning environment with some basic similarities but offering radically different affordances. The Second Life, system by Linden Lab is a persistent 3D world, or "metaverse". Users access the online system with a proprietary client and interact with content and other

"residents." Unique features include simple tools for constructing 3D objects and scripting tools for interactive content - including connectivity with external web-pages and internet resources.

Second Life is increasingly used in educational contexts. While educators have to deal with limitations in Second Life, it is a quickly evolving technology and remains the most accessible of these newer virtual worlds. According to the recently appointed Chief Executive Officer of Second Life Mark Kingdon, one goal is to continue with innovation as more competition increases from new virtual worlds[26]. Some early research suggests that virtual worlds, and in particular Second Life, may provide pedagogical advantages for specific learning styles and learner groups and for particular subject areas [8, 27, 30, 31]. One group of researchers report that Second Life is an educational game and that it should be both informative and engaging [30,32]. They conducted research on virtual world games, in particular Dollar Bay, which is intended to teach retailing principles and practices to students playing the game, it is meant to assist them in acquiring both real life skills and concepts as well. The rationale then is that this learning is subsequently transferred to real life situations [30]. One study investigated whether learning can be improved through interaction with a virtual environment. Although the results were inconclusive in this complex issue, they suggest that more investigation might be beneficial in increasing the body of research knowledge [27]. Another group of researchers also suggest that "there are limitless possibilities for educational research and evaluation within the Second Life environment" [29]. Another type of virtual world is SimCity which Devich (2008)[12] suggests as an example of a strategy game. This means that a player needs to win and to do so, must acquire a level of skill.

Teaching and learning activities in virtual worlds help pupils develop competences that can be used in other contexts as well be it school projects and/or their future work life [9]. It is particularly interesting because it was conducted with a group of pre-service educational students on the Teen Grid in Second Life as part of a final year course in educational technology. They completed a problem-based learning experience in Second Life that allowed the students to explore the virtual world as well as develop a course that could be taught to a high school class. It is suggested that manipulation and problem-solving skills that the students evolve when learning basic Second Life related skills are transferable to other activities. Other results referenced are: improved communication, teamwork, creativity and leadership skills. Some activities carried out focused on: Cyber Bullying; Going on excursions; Mapping; Scavenger Hunts; Role Playing; Communication Skills;

Children creating ideal classrooms, etc., with mixed degrees of success[9]

Kenneth Lim presents six lenses through which curricular interventions might be analyzed and evaluated, and suggests that learning goals in VW's can be maximized only if there is a mutually respectful relationship between school managers, teachers and developers informed by these Six Learnings. The author stresses the importance of understanding that no single in-world curriculum should be designed to meet all Six Learnings but rather focus on one or two of them.

The six lenses are as follows: Learning by exploring; Learning by collaborating; Learning by being; Learning by building; Learning by championing; Learning by expressing[18]. It might be prudent that we think in these six concepts when discussing what sorts of activities we should aim to support on TSL rather than thinking in Math Labs, Physics Labs, etc., and thus play to the strengths of the medium. This, of course, will call for a rethinking of the teacher's role. In this regard it is important to note that VW's in education should be seen as a supplement to, rather than a substitution for, traditional classroom teaching.

De Freitas and Neumann identify several new emerging challenges for teaching practitioners that will have wide implications for lesson planning and the relationship between teachers and their pupils. Most notably: The change of tutorial roles towards roles such as tutor-practitioner/practitioner-mentor; The emergence and use of tools that can give the opportunity to adapt and author different scenarios and conditions for learning; The greater empowerment of the learner to the extent that they may be able to explore environments freely and have control over the tools and content development, production and sharing; New opportunities for learners to reflect upon structured(e.g. formal) and semi - structured(e.g. combining informal and formal) learning activities (with tutors and in peer-to-peer learning situations)[11].

There are a number of universities using Second Life both to teach curriculum and as an extra curricular tool. Often it is used to represent a real life environment; for example, at Harvard Law School they have created Harvard Extension School's moot court, where students learn about court proceedings in the Second Life environment [23].   Another example of this is the Theatron 3 Project[32] building of replicas of historical European theatres (The Higher Education Academy English Subject Centre, 2006). Others, such as the Open University-based Schome community, have focused on specific learner groups and on allowing students a high degree of creative freedom. Universities are also beginning to use virtual worlds to support research. In Australia,

Swinburne University's Centre for Advanced Internet Architectures is using Quake 3 to model statistics and monitor their computer network as part of the L3DGEWorld project. [35] Stanford University is teaching virtual world creation in their computer science programs as well as using them in research[33, 5] contends that the evolution of online virtual environments is enabling new types of social science research and can be viewed as virtual laboratories. Although Second Life is increasingly used, research in this area is relatively scarce so the research and evaluation opportunities are limitless [29].

As the saying goes, "pioneers are the people who catch arrows in their backs." To be a pioneering educator in Second Life is definitely a challenge, and academia in general sometimes discourages educators from exploring new teaching methodologies that appear a bit "out there." Here are 7 tips to help you be as successful as possible in using Second Life for real life education: Spend as much time as possible exploring Second Life; Talk to other educators who are currently using Second Life for real life education purposes; Come up with clear and measureable goals for your academic use of Second Life; Publish or Perish!; Remember that Second Life is a platform for a wide range of activities; Work at Unlearning.

Over the past two years, Bourgas Free University and partner schools in Bourgas participated in a project called AVATAR(Added Value of teAching in a virTual woRld), related to the use of virtual worlds in teaching. In connection with this research was done among 110 teachers from six European countries [2]. [Figure 4.]

Figure 4. Venues for virtual meetings and seminars

Bulgarian participants in the project AVATAR [2] are from Burgas, lead by associate professor dr. Maria Monova - Jeleva from Computer Center of Economics and Management Sciences of Burgas Free University. It is one of the international team project AVATAR and moderator for Bulgarian participants. Teachers are from three Burgas schools and subjects that are taught are mathematics, informatics,

physics and astronomy, chemistry and environment, history and civilization, philosophy and professional degree courses in construction and architecture.

From the beginning in Second Life my students start with creating an avatar and after we start with environment of Second Life. The students were all exposed to a new and innovative technology. None of them have ever used this virtual world before. The students used a problem –

based scenario approach to investigate Second Life and determine activities they could use in a high school setting. These activities were innovative, varied, and appear to be achievable in that they can be taught to high school students, perhaps using a closed environment, such as some games.

Before choosing the topic of my lesson in Second Life, let us first think about what gaming and science have in common. Games and science share with each other that they are both games, although only games seem so at first view. They share, as well, the fact that they are both centered on "model-based thinking," although only science seems so at first view. To see this let me juxtapose a quote from the physicist and physics educator David Hestenes and from the game designer Will Wright, designer of such games as Sim City, The Sims, and Spore: The basic principles of Newtonian mechanics can be interpreted as a system of rules defining a medley of modeling games. The common objective of these games is to develop validated models of physical phenomena. This is the starting point for a promising new approach to physics instruction in which students are taught from the beginning that in science "modeling is the name of the game".[16] If you look at any kid playing a game, what they do is they go up and they grab the controller and they start pushing buttons randomly. They observe the results. They start building a model in their head for how the buttons are mapped. Then they start trying to set high-level goals. They start building a more and more elaborate model in their head of the underlying simulation in our game. And they're doing it purely through the scientific method. They observe data. They craft and experiment and do interactions to test their experiment. They observe their results then they increase the resolution of their model. And that's pretty much exactly what the scientific method is. So I think any kid, almost inherently, knows that and recognizes it as such.

David Hestenes says that Newtonian mechanics is a "tool kit" or an "engine" or a "design theory" that tells us how to build models of things like force and motion, models which are "toy versions" of reality (things like computer simulations, balls and ramps, or diagrams). We can then play with the toy (manipulate it according to certain rules) and, thanks to

its resemblance to the real world, we can make some educated guesses about how the world works and why it does so. If we are wrong, we have still learned something and we can change the model accordingly and try again.[16]

My experience in Second Life was held with students from Vocational Secondary School of Mechanics and Electronics – Burgas. The title of the lesson is "Stars. Diagram of Hertzshprung - Russell " from the syllabus of Physics and Astronomy - X Class, Standard IV.5 and VI [3]. The motivation for choosing the topic arises from the fact that studying the Universe in a virtual environment, students learn how information can be organized in a new ways (in the virtual space of Second Life). This helps them to learn more about stars, their similarities, differences and layout. Training in virtual environments helps them to be scientists who study stars and detect patterns and classes associated with them. In the real world they could hardly have achieved the effect of this to "build" their own objects like stars.

All of the buildings, interactive demos and content shown were built in-world using the Second Life modelling and scripting tools which clearly shows the flexibility of the environment. Similar simulations are readily available online and are usually created using Java or Adobe Flash. However creating the demonstrations in Second Life has a number of advantages. The virtual environment is immersive and allows residents to "walk around" and interact with the demonstrations in 3D space, viewing them from any conceivable angle. It also facilitates collaborative working and learning as groups of students and educators can occupy the same virtual space together. In addition the tools to create these demonstrations are reasonably straightforward to use.

The role of the scientist is extremely important in learning in virtual environments. Building Hertzshprung - Russell Diagram in Second Life, students carry out three steps that are crucial in the learning process: acquiring knowledge, practicing skills and knowledge assessment. Knowledge of the stars - This is the first step in students' work - an understanding of facts and concepts for the stars and it starts with a survey of theory and writing. The second aspect in the learning process is research, "review of the practice" of this knowledge in a particular environment and Second Life is the best for this topic. (Standard IV.5) [3]. Practice (in the virtual world) - Once students have grasped the concept of the structure of stars they will have to show how to implement it. They build stars of all spectral classes, starting from the Sun. The construction of the Hertzshprung - Russell diagram is crucial for the development of this practical. (Standard VI) [3]. Students have received knowledge of the

stars have learned to build the concept of spectral classes and how to build a Hertzshprung – Russell diagram, confident in their abilities.

The most important steps of the implementation of lessons for stars are: Development of theoretical basis for the stars (Notecards); Visit the Planetarium of Second Life, Exploratorium and other places related to the topic; Establishment of the Sun and other stars as objects; Construction of various stars of all spectral classes; Building a model diagram Hertzsprung - Russell, [Figure 5]; Development of models for the evolution of stars. To illustrate this point we will look at the snapshots from some of the demonstrations/simulations currently available on the island of the project AVATAR.
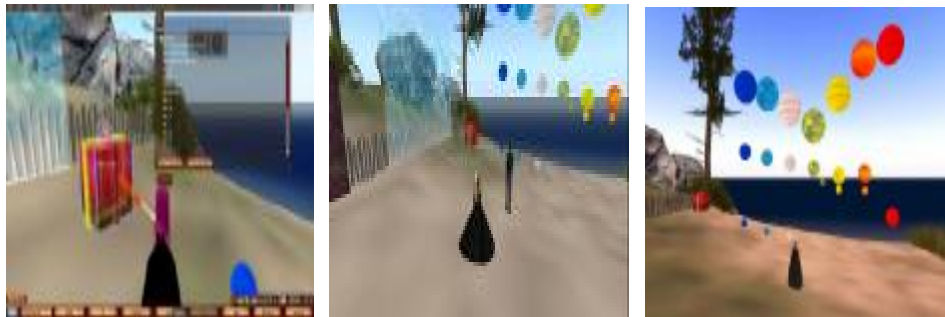

Figure 5. Stages of constructing models of star diagram Hertzsprung - Russell

The ability to take pictures, dance, drive, fly freely and pilot modern machines, as well as teleport from island to island, are just some of the pleasures of the students while teaching. [Figure 6]

Figure 6. Pictures whit student and teachers from events in Second life

In Second Life they have an opportunity for new types of training: through research, development, interaction, collaboration, journey through time and space study in another learning environment - the virtual environment of Second Life. Although designed specifically as an entertainment program, this environment is quickly recognized by teachers and researchers as an innovative, flexible and comfortable platform for communication and training. Through Second Life a student promote to love science and scientific culture of its inhabitants. We know mainly from our own experience and from interacting with other people that you acquire the best knowledge. All this is available in Second Life.

Overall, students felt that completing the activity in class made them more comfortable in exploring new technologies. They commented that they normally would not be able to experience these new technologies as well on their own. They also commented on how using Second Life in class opened their minds to new experiences. One of the students reported that Second Life "encourages students to develop their technical skills in second life." Other student commented that using new technologies in the future "we'll probably jump in there because we'll be a bit more familiar with how it works." While another student stated "I do see it as part of the 21st century learning – that's what it is." Through the completion of the task and subsequently the course, the pre-service education students achieved the overall goals of the course. Through the evidence gathered, including the focus group interviews, survey, work samples, and observation of the presentation, the conclusion can be drawn that the students achieved the goals of the course.

The virtual environment of Second Life is an integral part of the Global Outreach model. It acts as: An educational environment for teaching technology concepts; A tool for online collaboration; A platform for the development of a technological solution to a social problem.

Developed within the project AVATAR environment, methodology and resources allow for expansion of the traditional classroom education, providing tools and functionality for conducting virtual training in different subject areas with the possibility of applying different educational strategies and techniques while promoting active participation the students and their involvement in the learning process. The virtual world Second Life enhances the educational success of 21st century learners.

The virtual world is a very strong learning environment, which has a large set of tools for social interaction and innovation in education that encourage student participation. It has the ability to adapt and evolve according to the different needs of learners and can overcome the limitations of the situation imposed by the traditional classroom, where some tasks are difficult to be implemented for a variety of reasons (eg

equipment, location, etc.). Top 20 Educational Locations in Second Life we can see at [34].

Although basic research on the benefits of virtual world in terms of participation, control of the learner, educational standards and quality assurance are still in progress, the development of V-training in the next 5-10 years has the potential to radically change the way to learn and perceive education.

My name in Second Life is KALINAELI and I hope to meet you soon in Second Life!

Bibliography
1. Монова-Желева, М. Изграждане на знания и компетенции чрез виртуални светове. Информа принт, Бургас, 2011
2. Монова-Желева, М., Шурелова, Я. AVATAR – един курс за преподаване чрез виртуалния свят. Иновации в обучението и познавателното развитие, Образование и технологии, 2011,2
3. Учебна програма по физика и астрономия – Х клас, http://www.mon.bg/opencms/export/sites/mon/top_menu/general/educational _programs/10klas/fizika_10kl.pdf, [last 14.05.12].
4. AVATAR проект - http://www.avatarproject.eu, [last accessed 14.05.12].
5. Bainbridge, W. (2007) The scientific research potential of virtual worlds. *Science, 317*(5837), 472 - 476.)
6. Bartle, R. Designing virtual worlds, Indianapolis, IN: New Riders Publishing, 2003.
7. Bell, M. Toward a Definition of "Virtual Worlds", 2008, JVWR, Vol. 1(1).
8. Bradshaw, D. (2006) New practices in flexible learning: Virtual worlds - real learning! Pedagogical reflections. Retrieved 17th December, 2007, from http://virtualworlds.flexiblelearning.net.au/reports/VWRL_pedagog_reflect.pdf ; [last accessed 14.05.12].
9. Campbell, Chris. Learning in a different life, Journal of Virtual Worlds Research, 2009, Vol. 2, N1.
10. Craig S. D., Gholson B., Driscoll D. Animated Pedagogical Agents in Multimedia Educational Environments: Effects of Agent Properties, Picture Features, and Redundancy. Journal of Educational Psychology, 2004, Vol. 94.
11. De Freitas, S., Neumann, T. The use of 'exploratory learning' for supporting immersive learning in virtual environments,Computers& Education, 2009
12. Devisch, O. (2008) Should planners start playing computer games? Arguments from SimCity and Second Life. *Planning Theory & Practice, 9*(2), 209-226.
13. Fetscherin, M., & Lattemann, C. (2008) User acceptance of virtual worlds. *Journal of ElectronicCommerce Research, 9*(3), 231.
14. Freitas, S. de, Serious Virtual Worlds www.jisc.ac.uk/.../seriousvirtualworldsv1.pdf, 2008, [last accessed 14.05.12].
15. Gartner Group. (2007, 24th April) *Gartner says 80 percent of active Internet users will have a"Second Life" in the virtual world by the end of 2011.* [Press

Release] Stamford: Garner Research, from http://www.gartner.com/it/page. jsp?id=503861, [last 14.05.12].

16. Hestenes, D. (1992) Modeling games in the Newtonian world. *American Journal of Physics*,60, p. 732-748.

17. Kim, A.J., Community Building on the Web, 2000.

18. Lim, Kenneth. (2009) Editor's Corner: The Six learnings of Second Life: *A framework for designing curricular interventions in-world*, Journal of Virtual Worlds Research, Vol. 2. No. 1, Pedagogy, Education and Innovation in 3-D Virtual Worlds

19. Linden Lab. http://lindenlab.com/, [last accessed 14.05.12].

20. Linden Research. (2007) What is Second Life? Retrieved 19th December, 2007, from http://secondlife.com/whatis/

21. Linden Research (2008, 18th September, 2008). What is Second Life? Economic statistics. Retrieved 21st September, 2008, from http://secondlife. com/ whatis/economy_stats.php

22. Livingstone, D., & Kemp, J. (2008) Integrating web-based and 3D learning environments: Second Life meets Moodle. *Upgrade, IX*(3), 8 - 14.

23. Nolan, R. B. (2006) At law school, 'Second Life', in the cards, and the course catalogue. *TheHarvard Crimson: Online Edition* Retrieved 19th December, 2007, from http://www.thecrimson.com/article.aspx?ref=514500, [last accessed 14.05.12].

24. Pasolova-Førland, E., Sourin, A. and Sourina, O. (2006) Cybercampuses: design issues and future directions, *Visual Comput*, 22(12), 1015-1028, http://www.springerlink.com/content/u8188q1344331310, [last accessed 14.05.12].

25. Philip Rosedale - Virtual Worlds and Second Life, http://www.youtube. com/watch?v=J3LFqX6YNY0, [last accessed 14.05.12].

26. Prokofy, N. (2008). Interview: Linden Lab CEO Mark Kingdon. *Virtual Worlds News.* Retrieved 2nd May, 2008, from http://www.virtualworldsnews.com/ 2008/04/interviewlinde.html, [last accessed 14.05.12].

27. Roussou, M., Oliver, M. & Slater, M. (2006) The virtual playground: an educational virtual reality environment for evaluating interactivity and conceptual learning. *Virtual Reality,10*(3-4), 227.;

28. Second Life http://secondlife.com, [last accessed 14.05.12].

29. Sherman, G. & Tillies, R. (2007) *Educational research and evaluation data collectionprocedures in Second Life.* Paper presented at the World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education 2007, Quebec, Canada.

30. Slator, B. M., Chaput, H., Cosmano, R., Dischinger, B., Imdieke, C. & Vender, B. (2005) A multi-user desktop virtual environment for teaching shop-keeping to children. *Virtual Reality, 9*(1), 49. The Higher Education Academy English Subject Centre (2006, 11th December, 2007).

31. The Schome Community (2007). The schome-NAGTY teen second life pilot final report: A summary of key findings and lessons learnt Retrieved December 17, 2007, from http://kn.open.ac.uk/public/getfile.cfm? documentfileid=11344, [last accessed 14.05.12].

32. Theatron 3 - Educational undertakings in Second Life Retrieved 17th December, 2007, from http://www.english.heacademy.ac.uk/explore/projects/ archive/technology/tech23.php, [last accessed 14.05.12].
33. Timmer, J. (2007) Researchers hope to enable crowdsourcing of virtual worlds. *Ars Technica:The Art of Technology* Retrieved 30th March, 2008, from http://arstechnica.com/news.ars/post/20071230-researchers-hope-to-enable-crowdsourcingof-virtual-worlds.html, [last accessed 14.05.12].
34. Top 20 Educational Locations in Second Life, from http://www.simteach. com/wiki/index.php?title=Top_20_Educational_Locations_in_Second_Life, [last accessed 14.05.12].
35. Turner, A. (2008) Quake gives 3-D control. *Tech: Case Studies and Profiles.* Retrieved 26th April, 2008, from http://www.theage.com.au/news/biztech/ quake-gives-3dcontrol/2008/04/21/1208742844690.html, [last accessed 14.05.12].
36. Virtual Environment Info Group, Comparison table of virtual worlds, 2007, http://www.virtualenvironments.info/virtual-worlds-comparison-chart, [last accessed 14.05.12].

Author:
Yanka Georgieva Shurelova
Vocational Secondary School of Mechanics and Electronics – Burgas
Tel: 0888261295
E-mail: yana_shurelova@abv.bg

336

SCRATCH GAMES FOR PRESCHOOL STUDENTS

Veselina Spasova, Galina Momcheva

Abstract: The influence of using games in preschool training is quite important. A number of games from different categories are developed and are at beta testing level now. The action plan for developing Scratch games for pre-schoolers under the open educational resources ideas is prepared.

Keywords: games, preschool students, Scratch, OER

Today the society tries to reach the balance between digital natives and digital immigrants. We all are acquainted of such examples how little kids are playing computer games and have no difficulties in using a CD, opening a particular website or switching on the speakers.

It is clear for educators for ages that the game as a type of activity is social in its reasons, origin, content, structure and functions. It imitates the relationship between the player and the surrounding world, above all the social reality. To call an activity a game, it must possess the main structural elements of the game: an imaginary situation (game model); role; rules; game actions.

Educational games have to be incorporated in home and school environment.

The preschoolers are eager to learn by doing while playing a game. Educational games for children on this age are focused on physical, social, emotional and intellectual development.

Educational games for children often involve dramatic play, which helps to develop their social skills, as well as emotions as they are encouraged to express feelings with words and learn to share.

Offering games for intellectual development of preschoolers may involve understanding basic concepts, like distance, time, size, number or color. These classification games can help preschoolers develop reasoning skills.

The structure of computer games is built of three major functional components:

• The game environment is the subtotal of all the objects and relations in the game and the laws of their change. In other words, the game environment is the basis, the world in which the action of the game takes place.

• Interaction with players - this is the set of means by which players manipulate the game environment.

• Evaluation of game situation - these are the relations and the conditions that define the player's goal in a game. It includes a system of

awards (e.g. points) and penalties for game actions, as well as the description of the initial and the final game situation.

The most important of these components is the game environment. If it is well-chosen, the pace and the difficulty can be modified in different versions of the game by changing the other components.

When people play, they do not think much about the benefit from the outcome of the game. The game is an activity whose purpose is discovered in the process of playing. However, the lack of awareness of the practical importance of the game outcome can be accompanied by a subconscious result, which may be much more significant.

Chris Crawford described the term game along several dimensions such as art, entertainment, play, interaction, etc.

This approach may be summarized as an interactive, goal-oriented activity within which players (including virtual characters) can interfere with each other. An attempt to formalize the definition of game on the pillars of challenge, conflict and play came from Smed. These authors argue that the main components (the pillars) are linked together in a subtle way by the representation form (medium), by rules, by the goal definition, and by the absence or presence of opponents [8].

When a game is designed to teach or train during gameplay, the game is called an edutainment games due to the combination of education and entertainment in "one package". In edutainment games, the primary design focus is on the teaching part and game content is usually well-researched, designed around teaching principles or based on an actual curriculum.

The game activity applies the development function of the computer game by building the intellectual, emotional, volitional and ethical side of the individual in the different age periods:

At preschool age (3-6 years) children become like adults, but only in their minds, in their imagination. They seek to enter the adult world through imitation. The leading type of activity for this age is the game. It has double functions:

• Learning the rules and norms of life in society;
• Building awareness of their inner world [1].

The thinking of children at preschool age is intuitive. Logical thinking skills, used in the analysis of the relations between objects, are almost absent. The computer game can help children to establish the relations between objects, to combine and classify them. When creating games for children of this age group what must be considered is their great desire to play and to choose game situations, which require solving problems of the same type repeatedly and with increasing difficulty.

338

The game also shapes volition. It is associated primarily with the striving to some goal, i.e. purposefulness. The game teaches self-control, the ability to obey the rules of self-restriction. It expresses the inner strength of the person to manage their own experiences in significant situations.

The need to take decisions and to carry out specific actions for their implementation forms another important quality of the will: independence. It is expressed in taking initiative, in the absence of hesitation.

Game actions, whose result depends on the ability of players to evaluate and adjust their behaviour in view of productivity, are an incentive for self-organization, self-analysis and self-evaluation. This gives reason to point out the importance of the computer game in the self-perfection of the individual.

With the computer game people expand and deepen their knowledge of the surrounding world, they learn about the qualities and structures of objects, they study the world of human relationships by always taking an active part in them.

Depending on expected pedagogic result, there are several main applications of computer games:

- As a reward for the successful resolution of an academic task;
- As an incentive for serious work;
- As a means of modelling a research task.
- As a means of stimulating competition or cooperation.
- As a means of stimulating a certain type of thinking;
- To demonstrate the importance of and the relationship between various factors and situations;
- As a means of giving students an opportunity to practise a certain type of activity, which requires psycho-motor or cognitive skills.

The importance of educational computer games for the development function of the pedagogic process can be found primarily in the development of thinking.

Thinking is a process of acquiring summarised knowledge of reality. The computer game is an important tool for the development of the two main types of thinking - reproductive and creative, with variations in their relative share.

Reproductive thinking in education is connected with resolving educational tasks by reproducing samples. When the teaching goal asks for building this type of thinking, computer educational games must be of the training type using the following approach:

- Give the rule for solving a problem of a specific type;
- Analyse its application for a specific task;

- Assign other tasks of the same type in order to reinforce the algorithm for their solution.

Creative thinking is developed by educational computer games that intentionally place players in a situation of creative activity.

It is a known fact that the computer has the following two properties:

- Ability to present a wide range of examples, tasks and educational situations, where the learner can intervene and observe the results of their work.

- Ability to select and adapt the difficulty of tasks and situations and to give short instructions depending on the success or failure of learners, which are analysed by the programme.

The paradox in this new technological approach is that the subject of manipulation is not the brain of the learner (exposure to streams of audio-visual information), but that the learner manipulates with physical, chemical, and other objects, so that these actions cause the subtle changes in their brain.

According to the famous French psychologist Piaget, the roots of logic lie in motor actions and in the continuous taking pictures of (photographing) the results of these actions by the sensory systems and in fixing the relations between conditions - action - results.

The faster the technological approach is distributed in the education system, the closer we come to the school of the future, in which students will use laboratories, workshops and consultation centres, where they will listen a little and will ask a lot, but most of all will act.

The interaction determines the development of thinking in a specific field.

The computer revolution in education does not lie in transferring the contents of the classical book on screen. The creation of teaching software must be relatively independent from the existing educational content and its slow evolution [2].

Simulation models, built by modern computer equipment, play a major role in the development of sensorimotor logic as the basis of intuitive thinking, as well as of thinking as a whole, combining language and concepts on one hand, and images and actions on the other.

Here the methodology of the action approach is used. It is required for the development of specific frames of training programme and for defining the sequence of the tasks of a given class.

- The principle of support (example, directing visual image, technical devices)

- The principle of effect (the choice of situations must ensure a high percentage of success of trials, with the result visible, felt, or heard immediately.)

• The principle of isolation - parallel learning of several relations in the same action hinders their understanding and fixing in the brain, due to negative mutual influence (interference). Therefore simultaneous practice in more complex situations is acceptable only after the learner can distinguish clearly between individual relations.

• The principle of variation (a large number of examples of consolidation of the specific)

• The principle of gradation (learning with adequate speed)

• The principle of maturity (learning of relations and activities of a higher level after reaching the intuitive subconscious execution of the operations on a lower level.)

Some of the main features that make Scratch an extremely suitable environment for creating computer games for children are:

• Easy work in the environment combined with visual style of programming;

• Possibility for animating objects and adding sounds: music, speech, etc.

• Possibility for controlling various characters in 2D environment with mouse or keyboard.

• Options for sequencing selected actions (scripts), managed by events (the actions of the player);

• Possibility to change the theme of the computer game and its multiplication by changing the character or the costumes and preserving the logic;

• Possibility for publishing completed projects online.

The Scratch games that we are suggesting are explicit examples of open educational resources (OER) http://www.oercommons.org/. So teachers, educators or parents can modify, adapt or transform the games. It helps the kid from early ages to follow its own learning path guided by its desires and environment.

The idea of ScratchBG is to fulfill the original Scratch "Remix and share" idea with more reasoning practices like finding mistakes (tasks for programmers rubric [3]) without killing the creativity and the magic of randomness and to combine programming with activities like CS for fun and new types of CS unplugged tasks and ideas.

So the level of computational concepts deepens.

In a webinar in March 2012 Karen Brennan points out Computational Concepts.

These ideas are strongly supported by our team. The ScratchBG team is also involved in the idea to create own research actions for developping ScratchJr [9]. (a project led by Prof. Marina Bers from the

DevTech Research Group at Tufts University and by Prof. Mitch Resnick from the Lifelong Kindergarten group at the MIT Media Lab since 2011). It aims to develop and study the next generation of innovative technologies and curricular materials to support integrated STEM learning in early childhood education. It will develop, implement, and evaluate a new version of the Scratch programming language, ScratchJr, designed specifically for early childhood education.

The national standards for educational content for preschool groups in Bulgaria present the requirements for the development of quantitative notions by comparing objects through juxtaposition and superposing, classifying objects under given criteria, excluding objects from a group, comparing quantities in different groups, counting in straight and in reverse order, measurement, spatial notions referring to the identification of location - back/forth, left/right, up/down, basic geometrical forms (circle, triangle, square, rectangle).

The educational games designed here conform to the national standards for mathematical knowledge, skills and attitudes. They use plot elements from the films like 'Phineas and Ferb' etc. The individual elements of the above standards are applied in separate educational games, which makes them more convenient for the teacher/parent and correspond to the short attention span typical for the age.

Beside the adaptable parameters, which include change of characters, costumes (emotions) and backgrounds, the sound must also be changed, which asks for more time and effort.

We align a lot of our educational games to the national curriculum for preschool groups published at 2002. It states that these groups apply a new multi-ethnic model of preschool and primary school education. The social, emotional, motor, cognitive and artistic developments are interrelated. They are defined as educational priorities in the educational modules based on the varying kindergarten experience of the children.

The works of some authors are in the scope of our considerations. In his article Mr Kickmeier-Rust presents taxonomy of adaptive and adaptable digital educational games, including classifications, hierarchies, and rules. [7]
So we believe that the games that we've been created and want to develop as the OER idea also fostering collaboration and just-in-time learning as mentioned by Klopter. [4]

According to Klopfer some barriers to adoption of educational games are: curriculum requirements; attitudes (some parents and educators have negative attitudes about video games); logistics (Educators often find it difficult to integrate the play of game into the time structure of school day); support for teachers (Most teachers have little experience in integrating games into the classroom, and professional

342

development programs most often do not include support in this area); assessment (While games may be especially good at teaching higher order skills, these skills are not typically assessed in standardized exams); limited view (People often have a limited view of the variety of games available, which narrows expectations around the viability of games to engage students). [4]

But every child (person) at play is exercising freedom along five distinct axes: freedom to fail; freedom to experiment; freedom to fashion identities; freedom of effort; and freedom of interpretation. So, we have to take these benefits.

So in our case the level of teacher competency differs. One can only be a user of the readymade educational games another one can be a creator or modifier only. So putting educator in different possible roles by its choice for him/herself breaks in a natural way some of mentioned barriers.

Our experience begins with developing educational games for SEN children but soon the feedbacks gives us the idea that not only the SEN students' need such resources.

We have created a group of games under the idea of Montessori [6], and other following the pedagogical trends like CS Unplugged and JUMP Mathematics.

Now the experience continues with the practice with several kindergartens from Varna region. During our collaboration with educators for preschoolers and primary students we can make several conclusions that guide us to make the following action plan for developing the educational games and resources for preschoolers:

1. Enlarging the number and variety of educational games

Planning to create more educational resources that are described not only with the variety of content but also with the ability to organize and to exists with a self-organized online system (easy to access, publish and update the resources).

2. To add some extra tools like visual organizers;

3. Using/testing the new games and valuating feedbacks.

4. Optimize GUI of educational games after taking into consideration feedbacks.

5. Reaching high values in variability.

First implementations of games were static in their content, aims and resources. Such resources are easier to update but their effect is partial. That's why we experimented with putting more variability of the projects. So the level of their adaptability consists of the number of types of parameters that the teacher/parent can adapt, refresh or transform.

6. The forthcoming actions in our work is to teach preschool teachers to use Scratch in order to adapt educational games and for

those who are interested in it to continue teaching in advance level where they can create OER Scratch games alone or in teams.

7. Create flexible educators, not instructors

One of our basic ideas is to transform ICT in education from a jumping from one IT tool to another one to create flexible educators and also learners.

8. Preparing not only to use the resources but also to adapt from group of parents;

9. Interdisciplinary education

That's why a lot of interdisciplinary ideas have been discussed are leading in our work. It allows us to create projects classified as STEM projects, MAM (Music, Art, and Mathematics) projects and robotics for little kids.

In conclusion we wanted to implement LEGO bricks principles in creating new type of teacher – the flexible educator.

The educational resources package and instructions for teachers/parents are prepared for beta testing now and are available at http://www.scratch.bg/Pre-S_ver1.

References
1. Илиева, М. Компютърната игра, Народна просвета, 1989
2. Лалов, Б. Обучаващите програми, Народна просвета, 1990
3. Момчева, Г., З. Дженкова, Т. Момчева. Програмиране за всички. Scratch за начинаещи, Просвета, 2011
4. Eric Klopfer, Scot Osterweil, and Katie Salen. The Education Arcade, Massachusetts Institute of Technology, 2009
5. Jen Groff, Cathrin Howells and Sue Cranmer. The impact of console games in the classroom: Evidence from schools in Scotland, Futurelab 2010
6. Momcheva, G, V. Spasova. Games as a Challenge in Education, The 10th International Conference Information Technologies and Management, Riga, 2012
7. Michael D. Kickmeier-Rust. Talking Digital Educational Games (http://www.eightydays.eu/Paper/01%20Graz%20Workshop.pdf)
8. Smed, J. & Hakonen, H. (2003) Towards a Definition of a Computer Game. Technical Report, Nr. 553, Turku Centre for Computer Science, Turku
9. http://ase.tufts.edu/DevTech/ScratchJr/ScratchJrHome.asp - (The official site of the research team for ScratchJr)
10. http://mon.bg – The official site of the Ministry of Education Youth and Science
11. www.scratch.bg – The official site of ScratchBG group of professions and volunteers supporting visual programming for kids (lead by G. Momcheva)

NEW HORIZONS FOR COMPUTER SCIENCE IN EDUCATION

Natatasha Spiridonova

Abstract: During the last years the development of the Computer Science has attached a new social role to it. It becomes essential (due to the pressure of the contemporary social orders and expectations from the science), universal (due to the nature of the character of its internal cognitive system) and comprehensively applicable (due to the nature of the modern transfer of knowledge). This reveals new horizons and opportunities for its application.

In this scientific article we give arguments for the new place of the Computer Science in the university education. The problem is considered in a specific angle, which in some sense is external to the science itself. When we publicly update the problems of the Computer Science, it is natural for them to become the center of many educational programs. The attempt to introduce the Computer Science in all departments and scientific subjects in the university education is our professional obligation and responsibility. This would mean to comply in a literate way with the requirements of the modern reality.

The rapid developments in computer science in recent years have clad it in a new social role. It has become necessary (under the pressure of modern social demands and expectations of science), universal (due to its nature and the development of its inherent cognitive system) and applicable across fields (because of the nature of modern knowledge transfer). All of this has opened new horizons and created various application opportunities in the field of computer science.

This short paper (article) offers arguments in favor of computer science's new role in specialized university education and training.

The issue is approached from a specific perspective, which can be described, in a sense, as "extrascientific."

During the last third of the 20th century computer science has established itself as a separate field of knowledge developing its own methods and terminology. It is the science of collecting, transforming, transferring and storing information with arbitrary, including automatic, means. It studies information in terms of its structuredness, its quantitative aspects, its forms and ways of presentation. It also explores information processes as composition of the main information activities. As we know, com-position literally means "putting together." In other words, computer science is a systemic field of knowledge. Systems theory forms an organic part of it and it includes information culture in its entirety – the culture of information sources and recipients as well as the one expressed in information means and messages.

Undeniably computer science originates from and is closely connected to mathematics, linguistics and electronic engineering, but it has long since outgrown their confines. The use of scientific and technological developments, especially computers, in information processing, has led to establishing the branch of the science of information known as computer informatics. This branch is so large that in some universities, and even countries (mostly English-speaking ones[1]), the science of information is called "computer science." This is of course unfounded. As renowned Dutch computer science professor Edsger Dijkstra ironically noted, "Computer science is no more about computers than astronomy is about telescopes."

The advent of the world wide web – the internet – afforded a major impetus to computer science's development. All main functions and processes in the information age are increasingly organized on a network basis, which has been the tendency historically. Networks build a new social morphology for our communities. The proliferation of network logic has significantly altered the functioning and effects of the manufacturing, experience, cultural and, most importantly for the purposes of this paper, **scientific** processes. Network logic is a characteristic of the information technology paradigm.[2], [1] New opportunities for virtual connections construct a new type of sociality depending on users' tastes, interests and intellectual capabilities.

Computer science is a developing autonomous academic field open to other scientific spheres, rapidly adopting the opportunities they have to offer. Computer science is familiar with and applies new technologies within its sphere. That is not limited to the technological and information tools of the science and technology revolution (e.g. computers), but also extends to new social technologies developed in the humanitarian field. Computer science is the creator, herald and propagator of a new information culture integrating the effects and modern achievements of many other disciplines.

The internet, at least in theory, holds a great potential for every branch of science. It is almost as if it was created specifically for that purpose. Anyone who is connected (and now 36 % of Bulgarians are[2]) has an unlimited access to a huge and systematized information tank of verified themed knowledge. The internet is characterized by high speed,

---

[1] The term "informatics" is not used in English. It is, however, in use in German – Informatik, French – informatique, and other Romance languages – informatica.
[2] Data derived from Internet World Stats (http://www.internetworldstats.com/) as of late 2011, internet users in the USA comprise 78 % of the population, in Europe – 58 %, and in Bulgaria – 36 % and growing.

flexibility and adaptability to specific preferences. Through it one can produce science with a 24/7 focus group.

Some specialists consider online science to be the indisputable future of science as a whole. Even if this is an exaggeration (for instance, it is a fact that far from all scientific experiments can be prepared and carried out online), it is obvious that the possibility holds a number of advantages, such as the opportunity to circumvent publishers and editors, as one can skip the middleman. The information can be edited, linguistically modified and adjusted so as to meet the recipients' needs, which is impossible to do with the traditional tools. Space or time are no object. Communication with a larger audience is possible through arrays of messages – a continual electronic conference with certain and instantaneous feedback. The actual costs of any information activity (e.g. electronic instead of print magazines and publications) are significantly lower. There is greater variety of forms of expression which stimulates creative work. This list, rather extended and chunky on its face, can be expanded even further.

However, in order to make use of all those assets, one needs to be trained in computer science and able to navigate the internet freely, as well as possess a level of computer literacy that covers detailed knowledge of the information services the web has to offer. This kind of expert knowledge has to be acquired somehow and there is no better source for that than the system of education. Further, after its acquisition, the knowledge needs to be regularly updated, because the internet is an open, continuously developing system. All of this serves to reinforce the organic bond between education in all its forms and computer science.

Since the internet is changing everything, it is important that the related disadvantages be followed, together with any new risks, as well as the advantages. The least of those is that through use of the web, various scientific organizations, companies, institutions and even private citizens have taken on the role of publishers without having the necessary potential and competence. The biggest disadvantage is the lack of control mechanisms.[3] In between those two far ends are setbacks such as restricted audience – because there will always be people resisting the all-pervading electronic talkfest full of clamoring and anger, who would not participate – or the inability to tell the significant from the ephemeral, the truth from delusion, the deliberately sought outcome from authentic science. This leads us to believe that ambitions for the

---

[3] According to B. Mizova, G. Barakan "...Spending more time in cyberspace shows an emotional and mental instability..." [3]

universal establishment of electronic science are rushed, exaggerated and for the time being, unfounded.

All of those considerations do not in any way minimize the importance of the connection between computer science and every academic field, nor do they compromise the potential for positive development in that respect.

Computer science is a discipline of our time. Information transfer amounts in essence to transfer of meaning, or more accurately, chains of meaning. Fast and efficient exchange of large volumes of information is not only possible in this day and age, it is also organically necessary. There are three reasons for that. Firstly, without being in intensive contact with "the rest" who perform similar tasks and possibly already know how they can be done better, any activity can quickly become profaned. Secondly, the optimum professional work accentuates information both as a possibility and as an outcome. Contemporary specialist jobs are more communicative, internally, within their own structures as well as externally, with the objects and environment of their business. Thirdly, exchange of information simultaneously realizes the necessary diversification and cohesion in the contemporary world. Being enlightened means knowing and acknowledging differences. To attain that enlightened behavior, however, again we need systematic education in computer science.

Any branch of science and its respective academic field is in essence a targeted information activity. This what makes the connection between any academic field and computer science obvious and undeniable. Furthermore, it underlies the field's traditions at the root of its origination. The goals are informational (e.g. conception of new knowledge), the means are communicative (e.g. professional communication and academic discourse), and the intentions are also informational (e.g. all central and parallel efforts to disseminate scientific knowledge). In a sense, computer science knowledge is the be-all and end-all of scientific work. What this is means is that this knowledge must be made an integral part of university education.

It is obvious that information behavior is at the root of social cohesion. It does not cause the cohesion, but is how the masses of people become communities and societies. Communication is a factor of social alliance, whose purpose is that of the primary driving forces behind social development – people's material needs and their satisfaction. Communication serves the "official functions" of an indispensible and

essential condition (conditio sine qua non)[4] of joint human activity. In the USA in the 1990s 50% of employed citizens worked in the computer science field. Today this figure has risen to 60 %. With this kind of growth, computer science simply cannot be overestimated.

It should be noted, however, that all sources are not equal. As K. Nordström and J. Ridderstrale [4] point out, globalization means that "global stars capture all the attention. Why listen to the local bard or band when you can get a Madonna, Eagle-Eye Cherry or Pavarotti CD? Why use a local body builder in an action movie, when people are only interested in Sly Stallone, Arnie Schwarzenegger or Bruce Willis? Why watch a local soccer game when you can watch the best football team in Europe, Manchester United, playing Juventus on TV?" Why use the findings of our original, homebred scientific tradition if they cannot compare to the highest scientific gains on a global scale? The Yale information culture program holds a similar stance.

Computer science can alter language and the meaning of messages. Unlike the case with science, in the social realm "truth," "half-truth," and "lie" are very often relative terms and the main difference between them is the number of their repetitions in the public sphere.

On the other hand, computer science maintains that one of the most important qualities of the recipient in the act of information is his or her selective ability and the resulting potential to overcome communication thresholds. The modern person is deluged with information flows and it takes a number of special steps to choose a communicator, to find the information required and to share his or her opinion on the matter at hand. Each person has a certain information threshold beyond which information becomes redundant, tedious and dysfunctional. Information oversaturation is just as harmful as information deficit. It is impossible to inform everyone about everything. It is imperative that some kind of selection is realized on a systemic basis. This selective work, however, is in itself an additional burden on the professional communicator. Naturally, it is not error-proof, either. To avoid any errors and omissions, once again the answer is systematic understanding and knowledge in computer science.

It has to be added that information sources and recipients are not symmetrical in position, in significance, or in number. It is possible for one source to communicate with multiple recipients and for a number of sources to be directed at the same recipient. The audience and the

---

[4] Conditio sine qua non – a Latin phrase referring to an indispensable and essential action, condition, or ingredient

information stock are constantly vibrating – contracting and expanding. There is asymmetry in communication, but no position is privileged.

Some might consider the points proposed above a little far-fetched, maybe abstract and/or belonging to the realm of "high theory." This position, however, is not corroborated by the current scientific reality. It imperatively points to the fact that the necessary social and scientific conditions for computer science to be included as a compulsory course in all university programs for all academic fields are already present. Bulgarian educational practice already has a measure of positive experience in this respect.[5] This policy, if implemented, would yield positive results in several aspects. Firstly, students would benefit from it, because it would better equip them for their respective professional fields once they enter the job market. Secondly, existing problems (which are inherently inter-disciplinary) would find quicker resolution, because the transfer of knowledge between academic fields, necessary for tackling them, would be significantly facilitated. Thirdly, applying scientific findings in practice would be improved, as computer science helps foster better conditions for dialogue and interaction between science, business and social practice. In an absence of a relevant computer-science-based information culture it would be impossible to create and promote a knowledge-based economy.

In Weber's expression, education operates to meet "the demands of the say." Social realization of university programs in general and of computer science in particular is open to changes. The necessity of such changes is gradually and continuously demonstrated and acknowledged. The right road ahead is through objective and unpremeditated academic discussion aiming for consensus. This paper's intellectual goals are constructed along those very lines.

No illusion is being entertained here that mandating universal computer science education in university programs would be an easy or a smooth process. The academic community is an active and responsible entity and the only one with the relevant competence and legitimacy to initiate and implement the necessary changes. Overcoming the fear of modernizing university programs requires, all at the same time: firstly, removal of impediments, including the well-known stagnation of interested professionals in the field; secondly, openness to innovation, creativity and modern youth's energy; thirdly, strict conservation of high academic and professional standards as well as known and proven public interests.

---

[5] For example, compulsory courses in Marxist theory and in Russian in Bulgaria's Socialist period and currently, compulsory courses in English in university programs.

Integrating computer science in university education cannot be achieved through legal instruments and restrictions. The pursuit of universal solutions imposed "from above" is destined to fail. Before educational standards can be established, it is necessary to outline the framework, to carry out the respective long-term educational experiments (including through various curricula and schedules, educational and grading methods) and then have a competent scientific discussion, resulting in a consensus. This is where government bodies, employers of university graduates, and the educational institutions themselves would step in.

Contemporary educational sociologists unite around the idea that the social function of education in general and of academic disciplines in particular can be understood and appreciated in their entirety only if the processes in educational institutions are seen in interaction, while simultaneously accounting for the elements of social, economic, cultural and ideological reproduction within and through them. However, when issues pertaining to computer science are publicly brought into focus, it is natural that they become the centre of a number of academic programs.

It is our professional duty and responsibility to attempt to introduce computer science in all sections and disciplines of university education. It would mean that we meet the demands of contemporary reality in an enlightened manner.

References
1. Van Dijk, J. The Network Society. London: Saga Publications, 1999
2. Кастелс, М. Възходът на мрежовото общество: Информационната епоха: икономика, общество и култура. ЛИК, София, 2004
3. Мизова, Б., Баракан, Г. Социалнопсихологически особености на общуването в интернет. Сп. Психологически изследвания, кн. 1, 2003
4. Нордстрьом, К., Ридерстреле, Й. Funky Business: Бизнесът на бъдещето. ИнфоДар, София, 2003
5. http://www.internetworldstats.com/

Author:
Natasha Atanasova Spiridonova, PhD candidate at the PR department, South-West University "Neofit Rilski"
66 Ivan Michailov st,. 2700 Blagoevgrad, Bulgaria
e-mail: spiridonova@swu.bg

SOME APPROACHES FOR THE REALIZATION OF AN ADAPTIVE INTERACTIVE SYSTEM FOR DISTANCE LEARNING

Asya Stoyanova, Todorka Glushkova, Vanya Ivanova, Emil Doychev

Abstract: The article discusses some characteristics and methods of implementation of the interactive levels- feedback and control, productivity and creativity, communication and adaptation through the stages of modeling, and construction and real application of the first versions of the system for eLearning DeLC in the University of Plovdiv Paisii Hilendarski..

Key words: eLearning, DeLC, interactivity, SCORM, testing system

Recent trends in the development of education systems clearly identify the need to create standardized, adaptive, customized systems, which maintain a high level of interactivity and collaboration. An attempt to create such a system is the Distributed eLearning Center DeLC ([6] [7]) that is intended to provide personalized access to educational information and services anywhere and at any time in a synchronous and asynchronous mode. In the specialized literature there are different aspects to define the interactivity [10]. We will define it in the aspect of e-Learning as a dialog between users and the learning system in the training process and will view it at the following three main levels**:** ***Standard Experience,*** where the physical structure and hierarchy of the learning content remains unchanged and the student receives the entire information, notwithstanding if he knows it; ***Personal Experience,*** where the learning content doesn't change, but the hierarchy of the content changes and adapts to the user's behaviour and selections; ***Open Experience,*** that is connected with open and live systems with continuous engagement between the producers, users and the message.

The article discusses some characteristics and methods of implementation of these interactive levels and includes in them sub-levels, defined by Nathan Shedorff in [5] - feedback and control, productivity and creativity, communication and adaptation through the stages of modeling, and construction and real application of the first versions of the system. The authors share their experience in the development and implementation of personalized interactive services and resources used for students from the Faculty of Mathematics and Informatics at the University of Plovdiv Paisii Hilendarski.

A system for electronic and distance learning DeLC is developed by the Plovdiv University "Paisii Hilendarski"(PU) together with the Institute of Information Technologies (IIT), BAS- Bulgaria; the Telecommunication

Research Centre (TRC), the University of Limerick, Ireland; the Software Technology Research Laboratory (STRL), De Montfort University, Leicester, UK; the Software Technology Group (SWT), the Humboldt University, Berlin, Germany and the secondary school "Hristo Smirnenski", Brezovo, Bulgaria. In compliance with the main objectives of the project we establish a network of educational DeLC-portals that provide an adapted learning process to their customers and exchange learning resources and services.

The information society requires the application of new methods and approaches to the independent learning as well as to the classroom education of students. According to the characteristics of the education in Bulgarian schools DeLC-models and approaches will be applied, which focus primarily on the interactivity and adaptability of learning and some aspects of their application. The establishment of educational environment is based on an adapted nine-layer architecture model of the corporate portal of a Delphi group. For a particular realization of the educational portal the portal framework Liferay is used (http://liferay.com), into which there is implemented LMS of SCORM RTE. There are many services that support the training process in different subjects and raise the level of interactivity in learning [4].

The portal is a web site intended for a particular audience, which performs the processes of analyzing, collecting, processing and delivering of information and services according to the user's personalization. The front-end module of the system currently being developed is implemented by using a portal technology. This allows each user to have access to the services and resources according to their profile and security privileges. As it is well known one of the portal's major advantages is the possibility to integrate different components and/or services, which leads to increasing the functionality of the portal's building blocks. In this sense distributing the functions in different layers will increase its usability and effectiveness. There is couple of classifications and types of portals [9]. The most suitable for our goals is the corporate educational portal. Based on some research results our team decided to adopt the Delphi group's corporate educational portal model. On figure 1 is shown the architectural framework of the educational DELC portal 0.
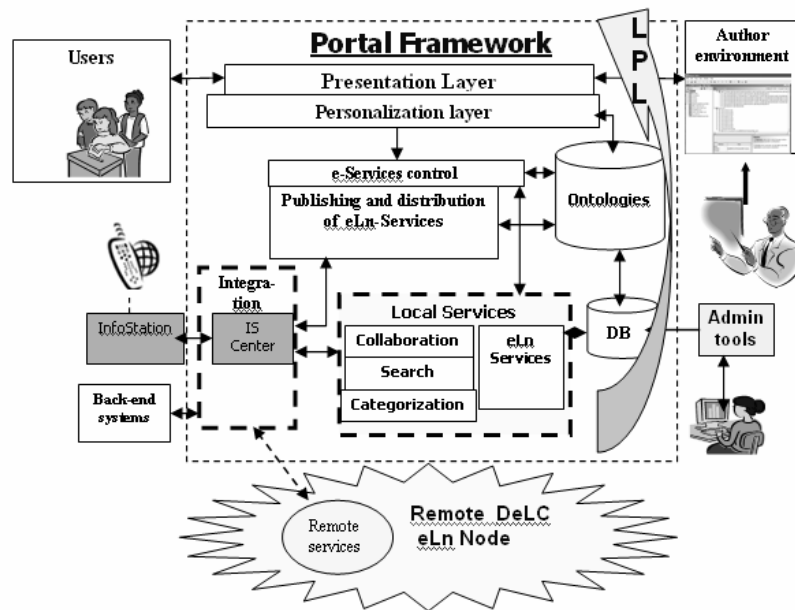
Fig. 1. Educational portal architecture framework

The framework contains 9 layers. Some of them contain sub-layers or are united in different groups.

- *The presentation layer* is the outermost layer, which ensures the interaction with the user. This layer gives the user access from one point to the whole system and allows the user to make personal settings for the user interface;

- *The personalization layer* ensures the interaction of the user with the portal. The Request analyzer is a sub-layer of this layer. Its role is to analyze the user's request and to deliver the requested services depending on the current domain, user profile and pedagogical profile of the student;

- *The e-Services control layer* is an important part of the portal. It ensures the access to the portal's services. It allows the users (according to their profile) to initiate and use different types of services related to the e-learning, or to interact with other users of these portals and other connected ones. This layer contains the DeLC navigator, which, according to the service type, forwards its management to the *Integration layer* or the particular LMS.

- *The publishing and distribution of the eLn-services layer* ensures the creation, authorization, addition, publication and sharing of

354

electronic educational resources and services. Here the problems are related to the correct formal description of the e-lectures and services and their publication according to the model rules in the expected period of time. This layer should ensure the operation of the SCORM RTE and the corresponding LMS.

– *The integration layer* – The DeLC e-services model, which is the basis for building the current prototype, includes 3 major groups of services – local, remote and external. This layer ensures that all services will be available and starts the corresponding run-time navigation and execution control. This is guaranteed by a suitable Gateway Management for each service type – Local gateway, Remote Gateway and Back-end Gateway. The Info Station Center (IS Center), developed together with our partners from the Limerick University, Ireland, is also located in this layer. The IS Center connects with a number of info-stations and thus gives an opportunity to use mobile services.

– *The collaboration layer* enables the usage of asynchronous (forums, off-line consultations) and synchronous (on-line consultations, chat-rooms) communications between users. This layer should ensure different levels of communications between users from various groups.

– *The search layer* provides tools for identification and access to a specific set of information resources – exams, lectures and so on. As the time goes by the collected educational information becomes quite large. This enforces the use of effective and intelligent mechanisms for detecting and delivering the most suitable information for a particular user.

– *The categorization layer* – different groups of users work with the portal and they use different educational resources and services. Therefore it is necessary to provide a mechanism for the categorization of the available information.

– *The loop process layer* – this layer differs from the others in the portal's architecture as it is not connected with specific aspects from the information flow management. It provides tools for a quick change of the workflow and management of the whole educational environment.

A characteristic feature of the discussed distributed portal architecture is the independence from the process of including and removing of particular services and educational resources. During the implementation of the different versions of the system these layers can be developed in varying degrees as they are continuously improved and expanded. Each layer ensures the achievement of a particular interactive level, according to the provided classification system. The ***Standard experience*** interactive level is connected mainly with the presentation

and personalization portal layers; the **Personal experience** interactive level can be realized in the integration, collaboration, search and categorization portal layers; the **Open experience** interactive level is used in the eService control for publishing and distribution of the e-Learning services and the loop process portal layers.

Standard experience

The first experience level is covered by the following subsystems:

• *WCM (web content management) system*, which is intended to support the process of creation, organization, publishing and management of the portal's static web content. This includes also strong access control mechanism, which contributes the portal's personalization facility (Fig. 2).
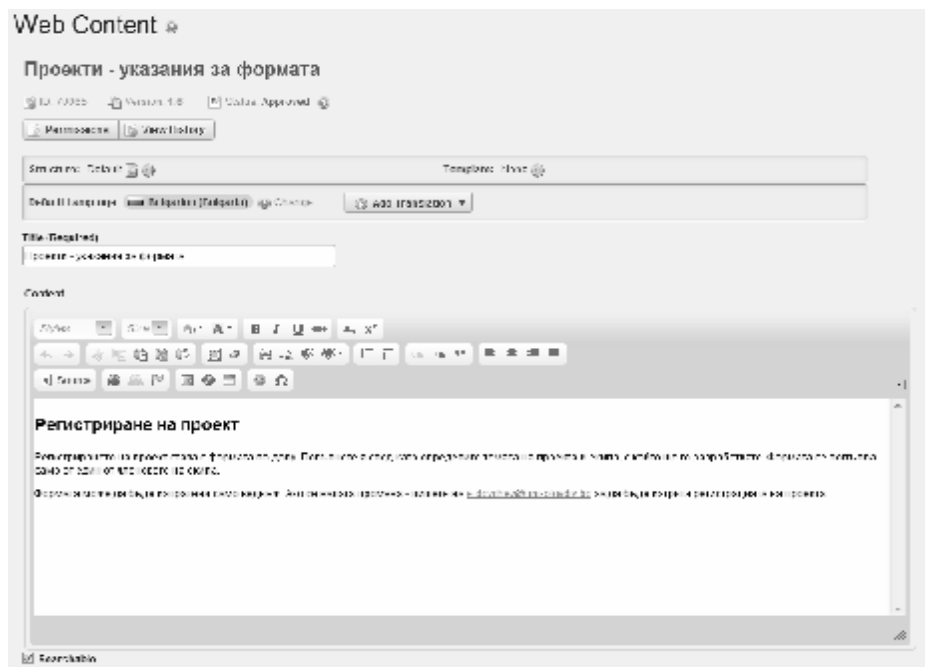

Fig. 2. Web content management

• *Documents and media library*. It is used to collect, organize and manage external documents, images and files in one place. The files can be organized in hierarchical folder structure and they are indexed by the portal's search tools. The content is made available to the users according to their rights. (Fig.3.)
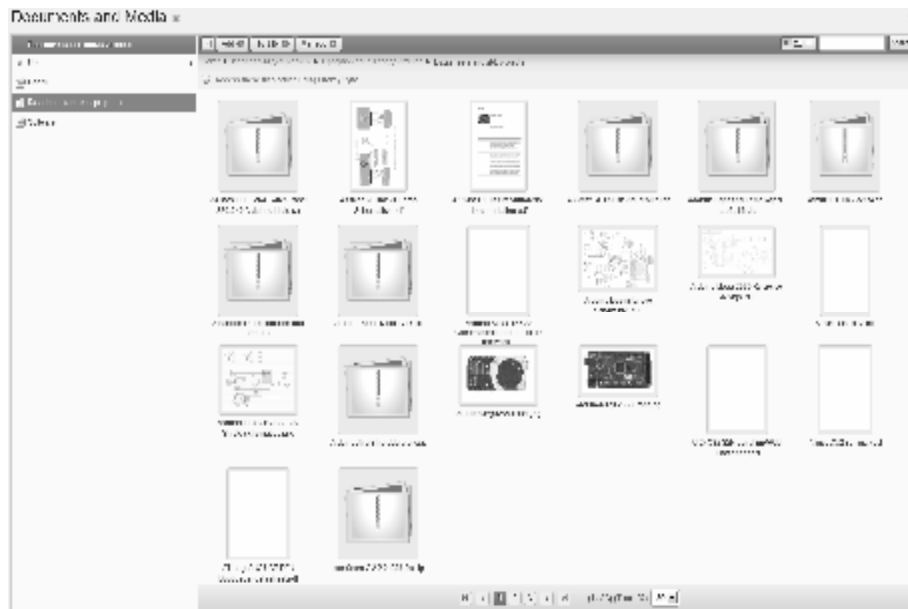
Fig.3. Documents and media library

Personal experience

The next experience level is covered by:

• SCORM Run-time engine. This upcoming subsystem is intended to make the SCORM content available to the portal's users. In contrast of the preceding interactive level the content presented here is not static. The content is delivered to the student according to their current state (related to the particular content) and the rules defined by the content creator.

• The test system is connected with this level. A test is defined as a series of questions, problems, or physical responses designed to determine knowledge, intelligence, or ability. All tests should serve a clear purpose and they have to be used "to render information to aid in making intelligent decisions about possible courses of action". Test authors have to be guided by principles in order to build good tests, which are professional and socially responsible. In DeLC we have followed the QTI standards in [11].

Traditionally, tests are used for one of five different purposes: achievement, aptitude, diagnosis, placement, and proficiency. However, testing can be used for other reasons as well such as to place learners into classes, to discover how much they have achieved, or to diagnose

difficulties that individual learners may have. Teachers also use tests to motivate learners to study [2].

In recent years the interest towards online testing as a means of assessing the knowledge and skills of students has increased enormously. Some of the main reasons for this are that computer tests have high scoring efficiency i.e. teachers can assess a large number of students in a limited time, which saves time and anxiety for correcting the tests and expenses for paper and photocopying. Besides, different students can be tested on different tests simultaneously.

A disadvantage of computer tests as compared to paper ones is in the area of navigation: it is easier for a test taker to move back and forth between questions on a paper test. However, all the test questions in DeLC are placed on a single sheet so scrolling up and down the page is easy. Text writing is also more convenient because of the editing possibilities of the testing system.

When a teacher composes a test in the Distributed eLearning Center DeLC, he or she can choose options from a Library with a set of questions and a set of templates.

The set of the questions includes the following types:

• multiple choice questions with one correct answer;

• multiple choice questions with more than one correct answers;

• strict matching (the number of questions corresponds exactly to the number of answers);

• non-strict matching (the number of questions does not correspond exactly to the number of answers; dichotomous items belong here too with the answers being only "True" or "False");

• open questions (these can require either a short answer consisting of a word or two in the case of which the questions can be scored automatically, while the long-answer ones must be evaluated by the teacher after the test has been submitted).

The templates are a set of predefined content typed in by the teacher. They are used to dynamically select a set of questions, based on criteria set by the teacher. A different test is generated for each student by using the templates included in the test.

The teacher can assign a certain number of points for each task and then change it in different tests in order to increase or decrease the weight of the question with regard to the relative importance of the topic for the particular group of students or the amount of time that has been spent on it in the classroom.

When a student completes their test and if there are no open questions in it they immediately receive their results. Currently a system for automatic assessment of all type of questions is under development

What students like about the tests in DeLC is above all the ability to see their test assessment online immediately (when long-answer open questions are included in the test students can see the number of points they have scored on the rest of the test items). Next, they appreciate the impossibility of anybody else to see their grades. They also use the resource library for easy sharing of materials and uploading resources. However, notwithstanding that each student works on a different test, they still try to cheat during the tests so teachers need to be vigilant.

On the other hand, lecturers like the dynamic nature of the tests and the flexible way for fine tuning the test assessment. The resource library contains a set of resources such as documents, images and archives, which can be used in the test creation.

The DeLC tests have a hierarchical organization and offer a strong and detailed access control, which can be set by the user (the student), their specialty, course and group.

An advantage of the tests in the Distributed eLearning Center is the optimization of the time for the whole exam. As the time counting is individual for all students, they do not normally all start and finish a test at the same time so as soon as a student submits their test and leaves the room, their free working place can be used immediately by another student.

Open experience

On the third interactive level (open experience) in DeLC we realized an intelligent environment that supports e-learning in software engineering – a refactoring Learning Environment (rLE) [8]. The main goal of the rLE is to analyze and assess the code, written by students in real time, as well as to recommend them changes in its structure, if needed, in order to improve its quality. The analysis and assessment are made by an intelligent assistant – a Refactoring Agent (RA) in compliance with the rules for refactoring for the programming language Java, defined in [1]. Depending on the refactoring method, which should be applied, the agent could react in three different ways:

1. To apply automatically the method after receiving confirmation from the student – automatic refactoring;

2. To display detailed instructions, explaining to the student where and how the particular refactoring method should be applied – refactoring proposal;

3. To ask the student additional questions in order to clarify the conditions and define the appropriate refactoring method – refactoring questionnaire.

Different reactions of the refactoring agent lead to different behaviour of the student. This behaviour is related to different decisions, which the student can make, when he/she develops the code. The student can be creative in his/her work i.e. the refactoring agent only assists the student in the refactoring process. It helps the student by offering advice in accordance with the refactoring rules, defined in the Knowledge Base.

In the cases when the refactoring method is relatively simple and the criteria for its execution are clear enough the agent could offer the student to realize the required changes automatically. In this situation the student has to make a decision:

1. Accept the offer from the agent, because he/she is not sure what to do with the code or simply likes the agent's proposal – in this situation the agent is an assistant that helps the student with the refactoring rules;

2. Not accept the agent's offer, because the idea, which lies behind the code, would be ruined – in this situation the student evinces creativity, and the agent only shows them some advice according to the refactoring rules in the Knowledge Base.

Often the refactoring criteria are clear but the execution of the particular method implies a significant change in the code or its structure. In these cases the recommended approach is for the agent to inform the student of the specific situation and propose to them detailed explanations about the possible improvements that could be made in the particular situation. The possible decisions of the student in this situation are:

1. To execute the proposed refactoring method – in this situation the agent guides the student by showing him/her a detailed list with steps for that particular refactoring method;

2. To refuse the proposed refactoring method – the idea of the current code of the student is different and the suggestion of the refactoring agent is not appropriate for it. In this situation the refactoring agent only assists the student by showing him/her a possible decision.

3. To use the proposed refactoring method but having modified it in the appropriate manner according to the particular case in the code and the particular goal – in this situation the student evinces creativity, because he/she uses current information about the refactoring method,

proposed by the refactoring agent, and applies additional knowledge to resolve the problem.

There are a lot of cases where the choice of applying one refactoring method or another is made on the basis of an almost one-type set of criteria where just a few of them differ from one another.

In the cases when some of the requirements for applying the refactoring methods are met and yet this is not sufficient to define explicitly the most appropriate one, the agent could "ask" the student one or more questions in order to clarify the concrete situation.

After the requirements become clear the agent defines the type of the situation again. It could then be brought to one of the above described types: automatic refactoring or a refactoring proposal.

When the agent is in the refactoring questionnaire situation, the student's response can be one of the following:

1. To answer the question that is asked by the refactoring agent. According to the answer of the student, the refactoring agent offers him/her a particular refactoring method that belongs to the refactoring proposal or an automatic refactoring. The response of the student depends on different possibilities, described in the previous situations.

2. Not to answer the question, asked by the refactoring agent - based on the asked question, the student can make the decision that the code is clear and there is no need to be refactored. In this situation the refactoring agent helps the student only by asking the question. The student is given the chance to think about the problem based on the content of the question. The student evinces creativity and can resolve the problem without a proposal of the refactoring agent.

A crucial role in the rLE is played by the Refactoring Agent that is the cornerstone in the proposed architecture. Different reactions of the agent lead to different behavior of the students. Deciding with the help of the Refactoring Agent which method of refactoring to use in the source code, the student can evince creativity. This makes the education of the student in refactoring more efficient and implements the "Learning by doing" strategy. The refactoring agent is an assistant that guides the student in refactoring and at the same time it gives the student an opportunity to make decisions alone and be creative while developing the code. The interaction between the refactoring agent and the student is a main part of the agent's activity, because this motivates the student to make decisions by themselves [4].

Conclusions

Contemporary trends in the electronic and distance learning require the realization of ever higher levels of interactivity and adaptability to the personal characteristics, knowledge and experience of each student. We are continuing to work on the implementation of an agent-oriented version of the e-learning system, as well as on the realization of scenarios related to adaptability in mobile learning. We have considered that the SCORM standard allows the creation of multiuse and personalized eLearning materials. The application of parameterized templates, scenarios and models for the creation of electronic tests and study materials in the second interactive level will increase the adaptability and interactivity of the system leading to an increase in its usability and effectiveness in training students in mixed and distance learning. The team is elaborating a model for the management of the Open experience interactive level by using of ITL and polices.

Acknowledgement

References

1. Fowler, M. Refactoring: Improving the Design of Existing Programs, Addison-Wesley, 1999
2. Fulcher, G. Practical Language Testing, 2010
3. Glushkova, T. S.Stojanov, M. Trendafilova, G. Cholakov. Adaptation of DeLC system for e-Learning in Secondary School, In Proc. Of International conference on Computer Systems and Technologies – CompSysTech'2005, Varna, Bulgaria 16-17 June, IV.15.1-15.6, ISBN 954-9641-38-4, http://ecet.ecs.ru.acad.bg/cst05/docs/sp/sIV/IV.15.pdf
4. Glushkova, T., A. Stojanova. Interaction and adaptation to the specificity of the subject domains in the system for e-Learning and distance training DeLC, In: Proc. of the International Conference "Informatics in the Scientific Knowledge 2008", 26-28 June 2008, Varna, pp.205-216, 2008. ISSN 1313-4345, ISBN 10:954-715-303-X, ISBN-13:978-954-715-303-5
5. Shedorff, N. Information Interaction Design: A Unified Field theory of design. In Jacobsen, R. (Ed.), Information Design, pp267-292. Cambridge, MA:MIT Press,1999.
6. Stojanov, S., I. Ganchev, I. Popchev, M. O'Droma. From CBT to e-Learning, Information Technologues and control, No 4/2005, Year III, pp. 2-10, ISSN 1312-2622.
7. Stojanov, S., I. Gancev, M. O'Droma, E. Dojchev, G. Cholakov. Agent oriented Distributed E-learning Center, International Conference, Varna, 2005, ISBN 3-929757-76-1, 1-st publ.

8. Stoyanov, S., A. Stoyanova-Doycheva, I. Popchev, M. Sandalski. ReLE - A Refactoring Supporting Tool, Compt. rend. Acad. bulg. Sci., 2011, Vol. 64, No. 7, pp. 1017-1026,   ISSN 1310-1331, IF 2010=0.219
9. Terra, J., C. Gordon. Realizing the Promise of Corporate Portals, Elsevier, 2002, ISBN-0750675934
10. Web-based Intelligent E-learning Systems, Idea Group Inc (IGI), 2005, ISBN-159140729X, 388p.
11. http://www.imsglobal.org/question/, visited on May 2012

Authors:
1. PhD **Asya Stoyanova-Doycheva**, Department of Computer Systems, University of Plovidiv "Paisii Hilendarski", Plovdiv, Bulgaria, astoyanova@uni-plovdiv.bg
2. PhD **Todorka Atanasova Glushkova**, SOU Brezovo, 0878212950, todorka.glushkova@gmail.com
3. **Vanya Ivanova**, Department of Education in Mathematics, Informatics and Information Technologies, University of Plovdiv "Paisii Hilendarski", vantod@uni-plovdiv.bg
4. **Emil Doychev**, Department of Computer Systems, University of Plovidiv "Paisii Hilendarski", Plovdiv, Bulgaria, e.doychev@uni-plovdiv.bg

ASSESMENT OF THE ABILITIES OF THE STUDENTS TO SEARCH INFORMATION AND TO COMMUNICATE

Dobrinka Stoykova, Silvia Stancheva

Introduction: In this article we aim to present our experience in using IT in teaching students of the specialty "Library Management". In particular we will examine methods of searching for information on specific topics and methods of virtual communication of our students with students from other countries. In our article we will give examples from the course "Information services in libraries". Finally we share our experience for making assessment of student's abilities to search information and to communicate virtually.

Key words: IT, library management, information services, virtual communication

The education of the students in the specialty "Library Management" is a challenge for the teachers. We expect from the students to achieve several goals. The first major aim is to build in the future librarians durable skills, needed in their everyday work. No less important is to teach them how to communicate successfully with other librarians and users. They have to analyze both internal and external environment and make quick and effective decisions in the interests of the library users. The students must learn how to seek alternative solutions in various situations (routine or extreme). The students learn also to develop creative scenarios, needed when they organize various events in the libraries.

To achieve these goals we use a combination of teaching methods, classical or interactive. One of the well known classical methods is the lectures. Through lectures, students receive a solid basis for seminars, practical exercises and application of interactive methods. The contents and the length of the lectures are determined by the volume of material, topics and objectives set by the teacher. The same applies to the interactive methods

Most frequently used from the interactive methods are – analyzes of real situations trough case studies, role plays and project work. New in our practice is the use of virtual communication. That includes searching, finding and disseminating information to partners who are in other countries. As in the classical, as well in the interactive methods by the exercises we widely use information technology (COM tools). In addition, we teach the students to work in groups of two or three. In that way they acquire the skills to work as a team.

In this article *we aim to present our experience in using IT in teaching* students of the specialty "Library Management". In particular we will examine methods of searching for information on specific topics and methods of virtual communication of our students with students from other countries. In our article we will give examples from the course "Information services in libraries". Finally we share our experience for making assessment of student's abilities to search information and to communicate virtually.

The course "Information services in libraries" is divided into three modules. In the first one are presented the sources of reference and information services in libraries. The second module brings various methods of reference. In the third module we speak about information services in various types of libraries. The sources for reference work in the library can be grouped into the following categories – catalogs (electronic or traditional), reference books, the book-stock, bibliographic editions, periodicals, databases (local and Internet).

Each of these types of sources can be in printed or electronic form. This requires a comparison between the information value of each type and source form. As far as possible in the lecture we avoid the citation of concrete titles. Emphasis is placed on the use of each type of source in a reference work. As a method we use "case study". That means to give typical examples for the use of the type or source in the work of our or foreign libraries. We point at advantages and disadvantages of each source type. The chosen method allows defining the main areas of application of different information sources.

During the practical exercises we concretize the lecture's contents. We give the students a number of practical tasks, by which they have to use various sources. At the beginning they make an unaided search for the publications of an author in the *online catalogue*. Then under the guidance of the teacher they have to find documents on a particular keyword. Initially, the teacher demonstrates a keyword search. Then he gives the students a task for self-searching. Finally, the teacher explains what the students didn't found or shows other ways of searching in the online catalogs. He demonstrates the basic, advanced or complex search, a search through the dictionary. Finally the teacher gives the students new tasks for self-searching, which they explain during the next exercise.

When searching for information *in printed reference books* we lay stress on typical problems and the elements in the publications, which make the search easier and faster. After acquiring skills for search in

printed reference books the students make that in electronic versions of encyclopedias, dictionaries, in handbooks and manuals and publications with biographical information in publications of geographic information in statistical yearbooks. To make sure, that the students can search without help we give them as homework to find unaided biographical information about a celebrity, factual information about scientific discovery and folk custom. In the search they must use printed and electronic reference books and in addition make their own conclusions about the quality and quantity of information found. These individual papers have to be checked and evaluated.

Special attention is paid to strategies for information retrieval. They require planning of the time, information databases and online access. This means to develop a tactical design for each specific information request in order to reduce information uncertainty by a series of successive searches, reducing the information noise and loss of information.

The negotiation of the inquiry is the first major step in the search. It has a direct impact on the quality of information found and on such indicators as the accuracy and completeness. During the exercises we use *the role play* to master the techniques in conducting reference interviews.

In the next stages of the search it is important to determine the search strategy. This includes the analysis of the reference question, to identify the key words, select the source, with which they start the search and making the search itself. In the exercises *work on case studies* /cases/, by which students learn how to analyze questions, determine what preliminary information is given and what the librarian should find for the customer. In some exercises we give more complex practical tasks – to search in electronic sources, by which the search lead to a change of tactics in demand. It require from any request to move to the use of a dictionary and /or Boolean search.

The criteria for meaningful line in the simple search include an assessment of the meaning of the title and annotation. The methods of inquiry (request) include: keywords, hierarchical lists of topics; masking (root word) sensitive search, boolean search, filtering by temporal, language, format characters, images, addresses, links.

The search procedures include the subject of the inquiry, choice of appropriate method of search, choice of keywords (specific rather than general); crossing terms (to comply with linguistic standards, synonymous phrases, specific terminology), analysis of relevance,

366

clarification of citations (authors, titles, URL addresses, dates of creation, of citations).

Electronic resources that future managers will use in the library may contain factual or bibliographic information. It is important for the students to be able to distinguish the selection of type source that contains the desired from the user information. It is known that electronic bibliographic resources contain bibliographic descriptions of published or unpublished documents in the thematic scope of the base. Such are for example the electronic catalogs of libraries, electronic bibliographies or thematic reference lists. Another part of the electronic bibliographic resources are with annotations or abstracts. These are branch databases with abstracts of scientific manuscripts or local bibliographies.

Another type of electronic information sources contain mainly factual information.. Such are for example the electronic encyclopedias, electronic dictionaries, and guides, address books, manuals or full-text databases with articles and/or books. Electronic bookstores also provide factual information about books and/or magazines that are for sale. In them the future managers can find a variety of useful information and rankings such as the best or best-selling works in print, presentations of the award-winning books, information about authors, extracts from books.

Regarding the access to information both groups can be online or offline. Typical of online information resources is that it allows remote electronic access to the information. This means that the users can search for information from their workplace or home in a database stored in a computer memory in a library or in an information center, often located miles away from his village.

In the course of "information services in libraries," students must learn where to search for and find lists of links to web sites and how to use them. Another part of the electronic resources, used in libraries, are offline. The information in them can be saved in a database, made by staff of the Library or be on a portable media, such as diskettes, CD's, CD ROM, DVD. This are for example the electronic "Bulgarian encyclopedia A-Z" and other electronic encyclopedias, offered in our market. The current National Bibliography of Bulgaria also has an electronic version. It is important to train the students to use such electronic editions.

It is important for students to learn how to do quality search in the web sites. It is therefore recommended to make short characteristics of the website, in which electronic sources can be found – analysis of the content, judgment of the speed and the accuracy in finding the desired

information. Therefore, in the assigned tasks we require to search at least in two places and then to compare the information regarding the completeness, speed and accuracy of search.

Each student is looking individually for bibliographical information on a given by the teacher topic. This is a thematic search of the titles of books and articles in the electronic bibliographic resources. The thematic list has an application form, in which the parameters for the search are specified. It contains a list of about 30 titles, described bibliographically, a list of sources, that are used by the search and the opinion of the student about the quality of the electronic resources. By this task it is not allowed to work in groups. Each student must demonstrate that he can select titles by subject and is able to present them in a way that is convenient for the user. In the thematically search the students use electronic catalogs of the leading libraries, bookseller's websites or in virtual libraries. Another important source for searching on a topic is the electronic current national bibliography.

Students from the specialty "Library Management" need also *communication skills.* They are necessary for their future information and consultancy activities. Such skills are needed also by the preparation of information leaflets for library information resources. On this important aspect of the library management we pay attention both in lectures and in seminars. We use a discussion or a Brain storming to discuss what flyers they need to prepare for a university or for a public library.

Last, but not least, our students communicate virtually with students and teachers from other countries. The beginning was in 2004, when *the virtual project* "Collaboration across borders" was created. It was for the program MINERVA/SOCRATES. On the homepage of www.cabweb. net [1] are pointed his goals, expected results and the pedagogical framework. *The Projects goals* are – to collaborate with colleagues in different countries; to be ready to bring the experience of international collaboration in classroom /course/ module; to seek the right partners for a collaborative work; to look for ways to internationalise the curriculum. Bulgaria joined in this project in 2006 at the invitation of the "Stuttgart University of Media". During the years we took part mainly in the forum JILID. Our partners were students and professors from more than 15 different universities in Europe, USA, Australia or Latin America.

The virtual portal Cabweb.net is created on the own server, which is results of the partnership between higher educational institutions from Poland, the Netherlands, Spain, Germany and United Kingdom, for

368

practicing and learning about collaboration between staff and students, for mutual educational benefit; [2]

The stage of the project JILID Community is wider the participants are more than 15 countries from Europe, Asia, Latin America and USA;

The technology of communication includes the formulation of the topics from the leader of the project.

The program logistics is made from two coordinators, who communicate interactive with all participants I follow up the dead lines, and activate the dialog;

The communication function of Internet that are most used are: Email, Chat, News groups;

Evaluation and giving feedback on partner's web-sites (multimedia presentations, programming projects, information systems etc.) based on established criteria;

Practice subject oriented discussion in second/foreign language and developing communication skills in target language.

One of the expected results in the forum JILID was: to build up a network for collaboration of staff and students. Other goals were: to facilitate different types of on-line collaborative activities with emphasis on critical evaluation and reflection; to investigate the benefit students derive from collaboration with peer groups at other institutions in Europe and across the world; to explore what makes the network sustainable rather then ephemeral ; to research on factors affecting collaboration, such as emotions and language barriers. All this was achieved until the end of the project in 2008. In the next years the collaboration was not so active and with less participants. Nevertheless the main body continued to exchange ideas and to communicate virtually on the created platform. Most of the professors participated also on the international ECE Conferences, hold in Salford, UK.

It was important for the virtual collaboration to make guidelines and support for tutors on designing collaborative exercises. The leading role in the virtual exchange took the German Professor Frank Tissen. For every session he formulated several questions, which had to be answered during the virtual communication. He invited the participants from different universities and determined the dead lines. In 2011 we received an invitation from him to do *a collaborations and discussions* on a *specific topic* with the help of web communication tools such as newsgroups, chats and documents on the web. Within a period of two months we should do an intensive communication about the subject. In that virtual communication took part students and teachers from: Austria,

Bulgaria**,** Brazil, Marocco, Mexico, Schweden, Sri Lanka, Turkey and some people from Belgium.

The virtual communication was done on the bases of *the theory of the cultural memory and the theory on the collective memory*. According to it every culture has specific places, things, heroes and stories that are important for the culture and the people to identify with, as well to present the values of the culture. Fairy tales are an example, but also buildings and places. The Eiffel tower in Paris is an example. The collective cultural memory may include heroes, persons or sometimes animals – for example the dragon in Chinese culture. The idea is that these cultural identification objects are present in the minds of the people and are used in information design, such as advertising, posters etc.

It was recommended to make following steps by the collaboration:

1. **Step 1**: an informal discussion about who is who AND a short introduction by everyone

2. **Step 2**: The participants start thinking of stories, places and heroes that are typical for their nation /culture and present the results to the others

3. **Step 3**: Discussions about the presentations among the participants and groups: What does that mean? Why are those stories, places, heroes important for you and your culture? What are the values behind? How do members of your culture learn about it? Is it still present today? What is your personal opinion about it?

4. **Step 4**: The participants try to find out how and where those stories, places, heroes are still visible in their culture (advertising, newspapers, TV, posters, and other information design products) and present it to the others.

5. **Step 5**: Discussion on it and a short documentation of the results. (What is similar among cultures? What can we learned from it?)

6. **Step 6**: Online party

We started the virtual communication in October 2011. The first thing we done in Bulgaria were to determine such typical stories, places and persons, which are important in our collective cultural memory. Then we divided the class in teams of three persons. Every team choose one or two objects and began to search information about it. They also discussed the results: What does that mean? Why are those stories, places, heroes important for our culture? What are the values behind? The Bulgarian students hold the discussion in following topics: Traditions

and customs; Natural landmarks; Traditional foods and drinks; National Holidays; Historic places; National heroes.

As teachers we found out that it is very useful for the students *to work as a team*. Aside from any required technical proficiency, a wide variety of social skills are desirable for successful teamwork. One of the most important from them is – the ability of every member of the team to listen and to respect the ideas of other members of the group. For a team to work effectively it is essential team members to acquire communication skills and use effective communication channels between one another e.g. using email, viral communication, group meetings and so on. This will enable members of the team to work together and achieve the team's purpose and goals. Other skills are: to ask questions, to interact, to discuss the objectives, to exchange and to defend ideas. It is important to take share in a joint task. Everybody from the team is expected to participate with what he does the best. In that case in a group of three students all of them are discussing the idea for participation in the virtual communication. Than one or two of them can search information in printed and in electronic resources. The person with the best English will translate the text. The person with the best information skills will present the message to the foreign partners.

Communication environment and COM tools

Every group of participants in the online discussion may send a message to a heading, formulated from the leader of the project or open a new rubric. The visible part of the message contains about 200 words and a photo of the participant. Additional on the screen can be displayed information in the size of altogether 500 words. It is possible to add also a picture, a photo and another composition.

By the search were used the resources of different libraries and the possibilities, which give the Internet. We used sources, which are actual, reliable, competent and accurate. The students selected also appropriate illustrations to every message. A part of the founded information was in English, but another part was in Bulgarian and needed to be translated.

Problems by the using of Com tools
Ø   The huge number of participants made it very difficult to follow up the discussion;
Ø   The email box of the participants was full of important or meaningless messages, which make communication impossible;

Ø   The topics area was not enough wide – only 200 words could be seen in front of the page and additionally 300 words;

Ø   The topics were formulated too broad and some of the participants sent their answers not to the exact topic.

Assessment of the student's knowledge

The assessment of student's knowledge can be classical or modern. From the classical methods we use exams, tests. By these methods there is a danger of memorizing the material. Therefore we use them in limited cases.

We use in much larger extent by the assessment of the student's knowledge a combination of modern methods that allow us to see if the studied material is understood. It should be noted that these methods are preferred by the students. There are such methods as a teamwork, in which students share their obligations or a self study and systematization of assigned practical tasks. They like to present to other students the ways of using the electronic resources of a certain library. Very well taken from the students is the participation in role plays.

Undoubtedly, the self use of electronic bibliographic resources to produce a bibliographic product is the best demonstration of student's knowledge. By that reason the assessment of the bibliographic list is taking part in the final grade in this subject. The learned from the students is used in greater extend during the specialized practice, which ends with a graduation work. That practice is the best test for the student's knowledge and experience.

Problems with the assessment of the answers in the virtual discussion

In connection with the assessment of the answers in the virtual communication we studied some publications about the problem [3, 4, 5 and 6].

According to Biggs [3] the assessment of the student's participation in the online discussion must be seen as a creative work. The teacher has to estimate highly the thought of the student, not only the facts.

In her summary about assessment Frances Bell outlines 3 main approaches to the assessment of the participation of the students in an on-line discussion [5]. According to Kim & Beal, 2006, it is easier to

measure the numbers of the sanded messages, their length and the numbers of answers to them from other students.

According to the second approach the teacher may use four criteria by the assessment of the student's messages- how often the students takes part in the discussion, how high is the quality of the sanded text, if to the text is added citation or a picture and if the student has worked in a team. This method is successful, but need a lot of time. [5, 4].

According to the third approach the assessment depends of the opinion of the students. They are asked to evaluate the best two contributions in the group. To be able to do this the teacher gives them a detailed instruction, where is explained how to evaluate the text. This method was successful only in the cases when the students didn't know each other or are not friends [5, 6].

We ascertained that it is possible to measure the quantitative responses such as numbers of messages and length of contributions. Much harder was to measure the qualitative info. We have to decide whether the students have engaged with the topic and posted meaningful contributions to the others. The assessment was influenced mainly from the demonstrated ability of the student to search information. On the other side we paid attention also on the quality of the presentation. For example a message with an appropriate picture became higher assessment compared with a message without a picture. The dialogue between students from different countries has been especially encouraged and facilitated.

Conclusions about the virtual communication

ü   The information on the Bulgarian websites on the above topics wasn't enough concrete;

ü   We had to search additionally in printed media about the meaning of the chosen topics in different messages;

ü   We made a lot of useful contacts with tutors from different countries;

ü   We could promote our culture in the entire world.

ü   Our students had established curiosity and a truly interest to virtual interactive learning.

References
1. http://en.wikipedia.org/wiki/Cultural_memory
2. http://www.cabweb.net/portal/course/view.php?id=36

3. **Biggs,** John. Aligning reaching for constructing learning. The Higher Education Academy. PDF. 4p.
4. **Wozniak,** Helen. Online discussions: Improving the quality on the student experience. PDF document. 10 p. - presenting the 5 year experiment at the School of Applied Vision Sciences at the University in Sydney, Australia.
5. **Bell,** Frances. Discussion Summary on Assessment of Online Discussion, held on the HELP network at http://www.cabweb.net 23/01/2006- 5/02/2006. 3 p.
6. **Stueart**, Robert D. Library and information center management. - Westport: Libraries Unlimited, 2007, 492 p
7. **Whatley,** Janice. Assessments Components for the evaluation of an on-line presentation of a group work of webpage. 2 p. // Information systems Institute, University of Salford, UK. Information on HELP forum
8. **Wilson,** A. P. Library web sites. Creating online collections and services. - Chicago, ALA, 2004.

Author:
Associate Professor Dobrinka Stoykova,
State University of Library Studies and Information Technologies,
Sofia, Bulgaria,
+359 (0) 2 970 85 89,
 dobrinka_s@abv.bg

PhD Silvia Stancheva Stancheva,
State University of Library Studies and Information Technologies,
    Sofia, Bulgaria,
    +359 (0) 2 970 85 89,
     mikovska@mail.bg

MATHEMATICAL BASICS OF INFORMATICS

EVALUATION OF THE ADVANTAGES OF CALCULATING MATHEMATICAL EQUATIONS USING A HASH-BASED STRUCTURES MAPPING KEY TO VALUE

Metodi Dimitrov

Abstract: *In the software world, many applications use the same equation. This requires their constant recalculation, leading to loss of time and resources. This article presents the data from a study comparing the time needed to calculate mathematical equations with the time needed to find the results in array of pre-calculated values.*

Key words: *mathematical equation, calculation time, hash table*

INTRODUCTION

Each computer program uses equation and performs calculations. As applications somewhat recreate the real world it sooner or later, the equations used to calculate various things will be repeated. The longer the humanity uses the computers, the more likely the same equations to be reused in different applications.

Such reuse and recalculation of the same equation leads to performing the same work on computers all over again. On the other hand, the complexity of the equations and the time to calculate them are related. That means the time of calculation increases with complexity. This would lead to irrational use of resources, computing power and inefficient work of computers.

An interesting fact is that even now many applications use the same equations for calculation of different things. If two different applications calculate the position of two objects using physical engine, then even though the objects are different - the equations that calculates their behavior are identical.

In addition to reuse of an equation in various applications, reuse can be observed within single application. It is very likely certain conditions to being repeated again and again within the same application. In this case the calculation of equations that are used for modeling of processes or events would be unnecessary.

EVALUATION OF THE ADVANTAGES OF CALCULATING MATHEMATICAL EQUATIONS USING A HASH-BASED STRUCTURES

A study that helps to determine is it faster an equation to be calculated or estimated result to be found among a group of pre-calculated values of the equation was conducted.

Storing pre-calculated results

Each equation is estimated in advance. That way during the execution the program doesn't waste time for calculation.

If the equation to be calculated is:

$$y=x^2+4 \tag{1}$$

the value of y is pre- calculated for various values of x. The calculated results are stored in hash-based structure mapping key to value. The key structure is the equation and the value - the result of calculation. For example, if x equals 1, according to (1) the key is "1*1+4" and the value is the results of the calculation - "5".

That way the equation is calculated in advance with thousands or millions of values of x and then the results are stored in the above-mentioned structure. Subsequently, if necessary the equation to be estimated (solved) with specific values instead of wasting time for calculation the ready response can be found among the many pre-calculated ones.

For purposes of the particular study each equation is estimated one million times with different random input parameters. The obtained data are stored in structures (described below) mapping key to a value. That way each structure contains one million pre-calculated equations with different input parameters.

Computer languages and operating systems, covered by the study

The study covers computer language C++, C# and Java. Results of pre-calculated equations are stored in generic classes mapping key to value. Selected structures are implemented using a hash table, which reduces the search time. Structures used are:

- for C++ - hash_map
- for Java - HashMap
- for c# - Dictionary

Computers used in this study have the following characteristics:
- Operating System - 32 bit and 64 bit Windows 7
- RAM - 2 to 3 GB DDR2
- CPU:
    - Cores - dual-core processors

- CPU clock – 1.7GHz to 2,1 GHz
- Type:
  - AMD AthlonX2 DualCore
  - Intel Pentium Dual CPU

Studied equations

The study covers two standard equations of physics:

1. Equation to calculate the time required for a body to fall freely from a height h1 to a height h2. Graphical visualization is shown in Fig. 1.

$$t = \frac{ar\cos\sqrt{\dfrac{h_2}{h_1}} + \sqrt{\dfrac{h_2}{h_1}\left(1 - \dfrac{h_2}{h_1}\right)}}{\sqrt{2\mu}} h_1^{3/2} \text{, s} \qquad (2)$$

where   h1 is the height from which the body falls, m

h2 - height to which the body falls, m
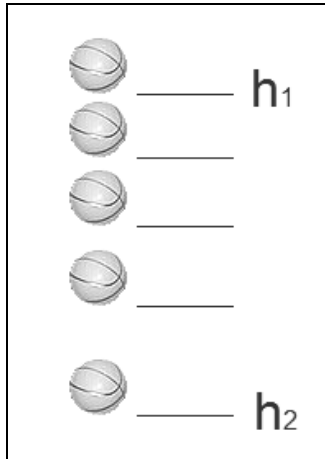
μ - a standard gravitational parameter, m^3/s^2

Fig. 1. Illustration of an equation to calculate the time to fall from a height of body height h1 to h2
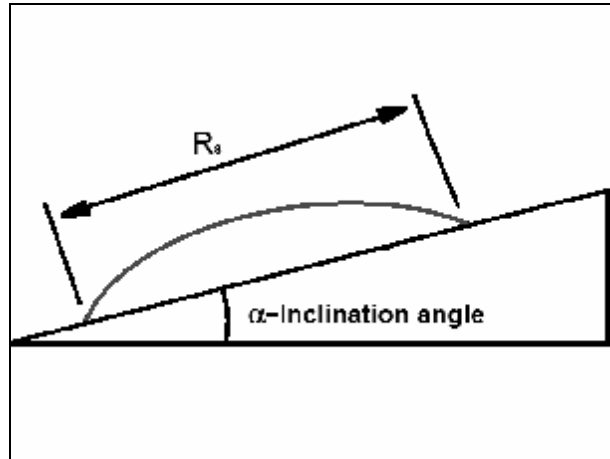
Fig. 2. Illustration of the equation calculating the distance at which an object falls on the sloping hill cast

2. Equation, calculating traveled distance of falling object, thrown on sloping hill. Graphical visualization is shown in Fig. 2.

$$R_s = \frac{2v^2 \cos^2 \theta}{g} \; \frac{\sin\theta}{\cos\theta} - \tan\alpha \; \sqrt{1+\tan^2\alpha} \, , m \qquad (3)$$

where v is the initial speed with which the object is been thrown, m / s
α - angle of the hill, º
Θ - angle of inclination of the object (against the horizon), º
g - gravity acceleration, m/s²
Standard mathematical operations, provided by each of the programming languages, are used for the calculation of the equations.

Study Results
Time needed to find the answer to the first equation - time in seconds that body falls from height h1 to a height h2 in both cases: calculation and searching the pre-calculated answer is presented on Fig. 3 and on Fig. 4. The results presented on fig.3 and fig. 4 are relate to one million iterations.
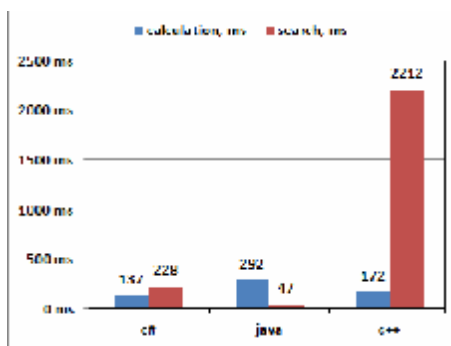

Fig. 3. Time required to find the answer to the equation 1 in both cases: calculation and searching the pre-calculated answer, rated to 1 million iterations
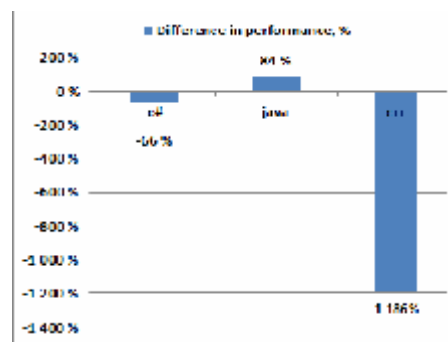
Fig. 4. Difference in rates between the time to calculate the equation 1 and a time to find the pre-calculated answer
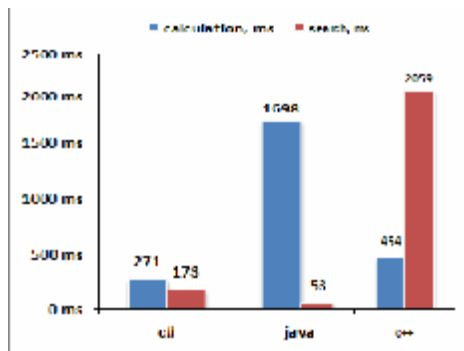
Fig. 5. Time required to find the answer to the equation 2 in both cases: calculation and searching the pre-calculated answer, rated to 1 million iterations
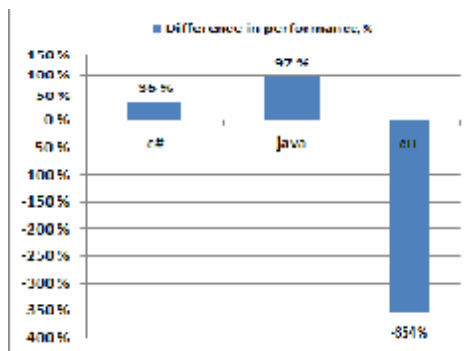
Fig. 6. Difference in rates between the time to calculate the equation 2 and a time to find the pre-calculated answer

Fig. 3 shows that there is an improvement in performance, when the answer is determinate by searching in the array of pre-calculated values, only in Java language. The calculation of one million equations took 292 ms and the search took only 47 ms Fig. 4 show that the time to find the answer in java was shortens by 84%. In other two languages (C++ and C#) calculating the answer is faster than searching the result in the array of pre-calculated values.

Time needed to find the answer to the second equation (calculating the distance at which an object thrown on the hill slope falls), again for both cases- calculation and searching the pre-calculated answer is presented in Fig. 5 and Fig. 6. The results presented on fig.5 and fig. 6 are relate to one million iterations.

Analyzing results for the second equation (fig. 5 and fig. 6) shows that there is improvement in performance, when answer is determinate by searching in the array of pre-calculated values in Java language (performance increased by 97%) and in C# language (performance increased by 36%).

Fig. 5 shows that the search time for C++, as in the case with first equation, is much greater than the calculation time. Comparing the data obtained in C++ with the data obtained in C # and Java allows assuming that a better implementation of the C++ container could improve the search time.

CONCLUSIONS AND FUTURE WORK

This study estimates time to find the answer of mathematical equations using a calculation and searching among a group of pre-calculated values.

Analyzed results show that depending on the equation in some cases it is faster to search for the pre-calculated value while in other cases it is faster the equation to be calculated.

In this line of thinking to make the search faster than computing, the equations should not be too small because the number of operations for search will exceed the number of operations for calculation.

Since the study covers only the operating system Windows, is appropriate in a further study be undertaken and other operating systems. It is in a later stage to develop an experimental platform and helps to study the possibilities for building applications, calculating mathematical formulas by searching the array of pre-calculated data.

References
1. Troelsen, A. Pro C# 2010 and the .NET 4 Platform, 5 edition, Apress, 2010.
2. Maurice, N., P. Wadler. Java Generics and Collections, O'Reilly Media, 2009.

About the author
Assistant Professor eng. **Metodi Dimitrov**, Department of Informatics and Information Technologies, University of Rousse, Phone: +359 82 888 470, E-mail: mdimitrov@uni-ruse.bg

ON THE UTILITY FUNCTIONS IN MATHEMATICAL MODEL OF EXCHANGE ECONOMICS

Zdravko Slavov

Abstract: *In this work we consider a mathematical model of a pure exchange economics and economic behavior of the economic agents in this model. It is easy to work with the preference relation of the economic agent, but when one wants to use calculus methods, it is easier to work with preferences that can be represented by a utility function. Here we will discuss the role of the utility function of the economic agent in exchange economics.*

Keywords: *exchange economics, mathematical model, preference, utility function.*

1. Introduction

Economics is a science that studies individuals economic behavior, economic phenomena, as well as how individual agents such as consumers, firms and government agencies, make trade-off choices that allocate limited resources among completing uses. Desires of peoples are unlimited, but resources are limited; therefore, individual agents must make trade-offs. We need economics to study this conflict and how there trade-offs are best made.

We will consider a mathematical model of a pure exchange economy $e(A, G, p, \{e^i\}_{i=1}^n, \{X_i\}_{i=1}^n, \{\mathbf{f}\}_{i=1}^n)$ with finite numbers of agents and goods. This economy is defined by a set A of economic agents, $|A| = n \geq 2$, $J_A = \{1, 2, ..., n\}$, a set G of exchange goods, $|G| = m \geq 2$, $J_G = \{1, 2, ..., m\}$, and a vector $p = (p_1, p_2, ..., p_m) \in R_{++}^m$ of prices of goods, $\sum_{i=1}^m p_i = 1$. Let each agent $a_i \in A$ have an endowment $e^i \in R_+^m$ (initial allocation) and a consumption set (set of bundles) $X_i \in R_+^m$, X_i is closed and convex, $e^i \in X_i$. The convexity of the consumption set means that every good is divisible and can be consumed by fraction units.

Note that a pure exchange economy is an economy in which there is no production. This is a special case of general economy. In this case, economic activities only consist of trading and consumption.

In addition, let denote the space of individual rational allocation (or rational space) by $\Sigma = \{x(x^1, x^2, ..., x^n) \in X_1 \times X_2 \times ... \times X_n \mid \sum_{i=1}^n x^i = \sum_{i=1}^n e^i\}$, where agent $a_i \in A$ owns of $x^i(x_1^i, x_2^i, ..., x_m^i) \in R_+^m$, a real number $x_j^i \geq 0$ shows the

quantity of $g_j \in G$ property of this agent. Clearly, we have that $e(e^1, e^2, ..., e^n) \in \Sigma$; therefore, the rational set Σ is nonempty. It is easy to show that the rational set Σ is compact and convex [3] [6].

As usually, the budget set of the economic agent $a_i \in A$ is given by

$$B_i(p, m_i) = \{x \in \Sigma \mid \langle p, x^i \rangle \le m_i\}, \qquad \text{where} \qquad m_i = \langle p, e^i \rangle \qquad \text{and}$$

$\langle p, x^i \rangle = \sum_{j=1}^{m} p_j x_j^i$ is the inner product of the price vector and consumption bundle.

Let each agent $a_i \in A$ have a binary weak preference relation $\underline{\mathbf{f}}_i$ defined on X_i, $\underline{\mathbf{f}}_i \subset X_i \times X_i$. A strict preference relation $\mathbf{f}_i$ is associated with $\underline{\mathbf{f}}_i$ as usual: $y \mathbf{f}_i x$ is equivalent to $y \underline{\mathbf{f}}_i x$ and not $x \underline{\mathbf{f}}_i y$ for $x, y \in X_i$.

From an economic point of view, the economic agent in the exchange economy $e(A, G, p, \{e^i\}_{i=1}^{n}, \{X_i\}_{i=1}^{n}, \{\underline{\mathbf{f}}_i\}_{i=1}^{n})$ is characterized by his initial allocation of the goods e^i, consumption set X_i, the budget set $B_i(p, m_i)$ and preference relation $\underline{\mathbf{f}}_i$. On the basic of these properties the economic agents exchange a finite number of perfectly divisible goods between each other.

2. The Properties of Preferences

To be precise, we introduce the following notations: for every two vectors $x, y \in R^m$, $x(x_1, x_2, ..., x_m) = y(y_1, y_2, ..., y_m)$ means $x_i = y_i$ for all $i \in J_G$, $x(x_1, x_2, ..., x_m) \geqq y(y_1, y_2, ..., y_m)$ means $x_i \ge y_i$ for all $i \in J_G$ (weakly component-wise order), $x(x_1, x_2, ..., x_m) > y(y_1, y_2, ..., y_m)$ means $x_i > y_i$ for all $i \in J_G$ (strictly component-wise order), and $x(x_1, x_2, ..., x_m) \ge y(y_1, y_2, ..., y_m)$ means $x_i \ge y_i$ for all $i \in J_G$ and $x_k > y_k$ for some $k \in J_G$ (component-wise order).

From an economic point of view for every $x, y \in X_i$, when we write $x \underline{\mathbf{f}}_i y$ we mean "the agent $a_i \in A$ thinks that the bundle $x \in X_i$ is at last as good as the bundle $y \in X_i$". Thus, we want the preference $\underline{\mathbf{f}}_i$ to order the set of bundles X_i. Therefore, we need to assume that any preference $\underline{\mathbf{f}}_i$ satisfies the following standard properties:

(1) Reflexive. For each $x \in X_i$, $x\underline{\mathbf{f}}_i x$.

(2) Transitive. For $x, y \in X_i$ then $x\underline{\mathbf{f}}_i z$.

(3) Complete. For every $x, y \in X_i$, either $x\underline{\mathbf{f}}_i y$ or $y\underline{\mathbf{f}}_i x$ or both.

(4) Continuous. For each $x \in X_i$, the upper contour set $P_i(x) = \{y \in X_i \mid y\underline{\mathbf{f}}_i x\}$ and the lower contour set $L_i(x) = \{y \in X_i \mid x\underline{\mathbf{f}}_i y\}$ are closed.

(5) Weak monotonic. If $x, y \in X_i$ and $x \geq y$, then $x\underline{\mathbf{f}}_i y$.

(6) Monotonic. If $x, y \in X_i$ and $x > y$, then $x\,\mathbf{f}_i\,y$.

(7) Strong monotonic. If $x, y \in X_i$ and $x \geq y$, then $x\,\mathbf{f}_i\,y$.

(8) Weak convexity. If $x, y \in X_i$, $y\underline{\mathbf{f}}_i x$ and $t \in [0,1]$, then $tx + (1-t)y\,\underline{\mathbf{f}}_i x$.

(9) Convexity. If $x, y \in X_i$, $y\,\mathbf{f}_i\,x$ and $t \in [0,1)$, then $tx + (1-t)y\,\mathbf{f}_i\,x$.

(10) Strict convexity. If $x, y \in X_i$, $y\underline{\mathbf{f}}_i x$ and $t \in (0,1)$, then $tx + (1-t)y\,\mathbf{f}_i\,x$.

3. Preferences and Utility Functions

Sometimes it is easier to work with the preference relation of the economic agent, but other times, especially when one wants to use calculus methods, it is easier to work with preferences that can be represented by a utility function.

Now we assume that every relation $\underline{\mathbf{f}}_i$ of $\{\underline{\mathbf{f}}\}_{i=1}^n$ is reflexive, transitive, complete and continuous on X_i. Thus each binary relation $\underline{\mathbf{f}}_i$ can be represented by a continuous utility function $u_i : X_i \to R$ such that for every $x, y \in X_i$, $x\underline{\mathbf{f}}_i y$ is equivalent to $u_i(x) \geq u_i(y)$ [1, Theorem 5.1] [3, Theorem 9].

Example 1. (Cobb-Douglas Utility Function) A utility function that is used frequently for illustrative and empirical purposes is the Cobb-Douglas utility function, $u(x_1, x_2, ..., x_m) = A x_1^{a_1} x_2^{a_2} ... x_m^{a_m}$ with $a_i > 0$ for $i = 1, 2, ..., m$ and $A > 0$. This utility function represents a preference ordering that is reflexive, transitive, complete, continuous, strong monotonic and strictly convex in R_{++}^m.

Example 2. (Linear Utility Function) A utility function that describes perfect substitution between goods is the linear utility function, $u(x_1, x_2,..., x_m) = a_1 x_1 + a_2 x_2 + ... + a_m x_m$ with $a_i \geq 0$ for all $i = 1, 2,..., m$ and $a_j > 0$ for some $j = 1, 2,..., m$. This utility function represents a preference ordering that is reflexive, transitive, complete, continuous, monotonic and convex in R_{++}^m.

Example 3. (Leontief Utility Function) A utility function that describes perfect complement between goods is the Leontief utility function, $u(x_1, x_2,..., x_m) = \min\{a_1 x_1, a_2 x_2,..., a_m x_m\}$ with $a_i \geq 0$ for all $i = 1, 2,..., m$ and $a_j > 0$ for some $j = 1, 2,..., m$. This utility function represents a preference ordering that is also reflexive, transitive, complete, continuous, monotonic and convex in R_{++}^m.

It is very important that not all preference orderings can be represented by utility functions.

The following statement shows the existence of a utility function when a preference ordering is strong monotonic.

Theorem 1 (Existence of a Utility Function). If the relation $\underline{\mathbf{f}}$ is reflexive, transitive, complete, continuous and strong monotonic on R_+^m, then there exists a continuous utility function $u : R_+^m \to R$ such that for every $x, y \in R_+^m$ if $x \geq y$, then $u(x) > u(y)$ (u represents $\underline{\mathbf{f}}$).

Proof. Let s be the vector in R_+^m consisting of all ones. Let also denote a new binary relation $\approx \in R_+^m \times R_+^m$ such that for every $x, y \in R_+^m$, $x \approx y$ is equivalent to $x\underline{\mathbf{f}}y$ and $y\underline{\mathbf{f}}x$. Then for a given vector $x \in R_+^m$ let $u(x)$ be that real number such that $x \approx u(x).s$.

Let $x \in R_+^m$, and denote $B(x) = \{t \in R \mid t.s\underline{\mathbf{f}}x\}$ and $W(x) = \{t \in R \mid x\underline{\mathbf{f}}t.s\}$. Then strong monotonicity of $\underline{\mathbf{f}}$ implies that $B(x)$ is nonempty and $W(x)$ is certainty nonempty since it is contains 0. Continuity of $\underline{\mathbf{f}}$ implies both sets are closed. Since the real line is connected, there is some $t_x \in R$ such that $x \approx t_x.s$.

Now let consider a function $u : R_+^m \to R$ such that $u(x) = t_x$.

First, we will prove that this function actually represents the preference $\underline{\mathbf{f}}$.

Let $x, y \in R_+^m$, $u(x) = t_x$ where $x \approx t_x.s$, $u(y) = t_y$ where $y \approx t_y.s$.

Then if $u(x) = t_x < t_y = u(y)$, strong monotonicity shows that $t_x.s \mathbf{p} \, t_y.s$, and transitivity shows that $x \approx t_x.s \, \mathbf{p} \, t_y.s \approx y$, i.e. $x \, \mathbf{p} \, y$.

Similarly, if $x \, \mathbf{p} \, y$, then $t_x.s \, \mathbf{p} \, t_y.s$. As a result we obtain $t_x < t_y$, i.e. $u(x) < u(y)$.

Second, we will prove the function u defined above is continuous on R_+^m.

Suppose $\{x_k\}_{k=1}^\infty \subset R_+^m$ is a sequence with $\lim_{k \to \infty} x_k = x_0$. We want to show that $\lim_{k \to \infty} u(x_k) = u(x_0)$. Let assume that $\lim_{k \to \infty} u(x_k) \neq u(x_0)$. Then we can find $e > 0$ and an infinite set of $\{k \,|\, k = 1,2,...\}$ such that $u(x_k) > u(x_0) + e$ or an infinite set of $\{k \,|\, k = 1,2,...\}$ such that $u(x_k) < u(x_0) - e$. Without loss of generality, let assume the first of these. This means that $x_k \approx u(x_k).s \, \mathbf{f} \, (u(x_0) + e).s \approx x_0 + e.s$. So by transitivity, $x_k \, \mathbf{f} \, x_0 + e.s$. But for a large k in our infinite set, $x_0 + e.s > x_k$, so we obtain $x_0 + e.s \, \mathbf{f} \, x_k$. This leads to a contradiction; therefore, the function u is continuous on R_+^m.

The theorem is proven.

Now, our attention will be focused on the preference $\underline{\mathbf{f}}$ and the utility function u.

Theorem 2 (Invariance of Utility Function to Monotonic Transforms). If the function u represents some preferences $\underline{\mathbf{f}}$ and $f : R \to R$ is strictly monotonic increasing function, then the function $f \, \mathbf{o} \, u$ will represent exactly the same preferences.

Proof. This is because for $x, y \in R_+^m$, $f(u(x)) \geq f(u(y))$ if and only if $u(x) \geq u(y)$.

The theorem is proven.

The invariance theorem is useful in many aspects in exchange economics. The utility function is a measure for the status quo of the economic agents.

Remark 1. The important properties of a preference ordering can be easily verified by examining utility function, see also Examples 1, 2 and 3. The properties are summarized in the following statement. Let $\underline{\mathbf{f}} \subset X \times X$ be represented by a utility function $u : X \to R$. Then:

(1) An ordering $\underline{\mathbf{f}}$ is strong monotonic on X if and only if u is strictly monotonic on X.

(2) An ordering $\underline{\mathbf{f}}$ is continuous on X if and only if u is continuous on X.

(3) An ordering $\underline{\mathbf{f}}$ is weakly convex on X if and only if u is quasi-concave on X.

(4) An ordering $\underline{\mathbf{f}}$ is strictly convex on X if and only if u is strictly quasi-concave on X.

Remark 2. Note that:

(1) A function u is quasi-concavity on X if and only if for $c \in R$ and $x, y \in X$, $u(x) \geq c$ and $u(y) \geq c$ implies that $u(tx + (1-t)y) \geq c$ for all $t \in (0;1)$.

(2) A function u is strictly quasi-concavity on X if and only if for $c \in R$ and $x, y \in X$, $u(x) \geq c$ and $u(y) \geq c$ implies that $u(tx + (1-t)y) > c$ for all $t \in (0;1)$.

Quasi-concavity of utility function plays a central role in exchange economics and optimization theory. Therefore, the research on quasi-concavity is one of the most important aspects of the optimization technique in mathematical model of exchange economics with perfectly divisible goods and quasi-concave utility function of the economic agents. For more information see [4], [5] and [6].

4. Utility Maximization and Optimal Choice

A fundamental hypothesis on individual behavior in modern economics in general and the consumer theory in particular is that a rational agent will always choose a most preferred bundle from the set of affordable alternatives.

In the basic problem of preference maximization, the set of affordable alternatives is the set of all bundles that satisfy the budget constraint. This problem can be written as:

maximize $u(x)$

subject to $x \in B(p, m_0)$,

where $p \in R_+^m$, $m_0 > 0$ and $B(p, m_0) = \{x \in X \mid \langle p, x \rangle \leq m_0\}$.

We know that there exists a solution to this optimization problem if the utility function u is continuous and the constraint set $B(p, m_0)$ is compact.

Remark 3. It is easy to see that the constraint set $B(p,m_0)$ is compact for $p \in R_{++}^m$. If some price is zero, then the consumer might want an infinite amount of the corresponding good; therefore, the constraint set $B(p,m_0)$ is not compact.

The utility maximization is a basic optimization problem in exchange economic theory. It is known that the optimization problem

maximize $u(x)$

subject to $x \in B(p,m_0)$

is equivalent to an optimization problem

maximize

$u(x)$ subject to $x \in B'(p,m_0)$,

where $B'(p,m_0) = \{x \in X \mid \langle p,x \rangle = m_0\}$.

The value of $x \in B'(p,m_0)$ that solve this problem is the consumer's demanded bundle, i.e. it expresses how much of each good the consumer desires at a given level of prices and income. In general, the optimal consumption is not unique. Denote by $x(p,m_0)$ the set of all utility maximizing consumption bundles, $x(p,m_0) \subset B'(p,m_0)$. This allows us to define a multifunction $x : R_+^m \times R_+ \to B(p,m_0)$ and it is called the consumer's demand multifunction. When there is a unique demanded bundle for each (p,m_0), $x(p,m_0)$ becomes a function and thus is called the consumer's demand function. We know that if the preference is strictly convex, then for each $(p,m_0) \in R_{++}^m \times R_+$ there is a unique bundle x that maximizes u on $B(p,m_0)$.

5. The Maximum Theorem

The Berge Maximum Theorem has become one of the most useful and powerful theorems in optimization theory and mathematical economics. The original variant of the Berge Maximum Theorem, shortly the Maximum Theorem, is as follows:

Theorem 3 [2] [7, Theorem 9.14]. Let $X \subset R^n$ and $Y \subset R^m$, $u : X \times Y \to R$ be a continuous function, and $D : X \Rightarrow Y$ be a compact-valued and continuous multifunction. Then, the function $m : X \to R$ defined by $m(x) = \max\{u(x,y) \mid y \in D(x)\}$ is continuous on X, and the multifunction $S : X \Rightarrow Y$ defined by $S(x) = \{y \in D(x) \mid u(x,y) = m(x)\}$ is compact-valued and upper semi-continuous on X.

The Maximum Theorem is often used in a special situation such that the multifunction D is convex-valued and the function u is quasi-concave or concave in its second variable in addition to the hypotheses of Theorem 3.

Now, we give a presentation of the classical variant of the Maximum Theorem.

Theorem 4 [7, Theorem 9.17 and Corollary 9.20]. Let $X \subset R^n$ and $Y \subset R^m$, $u : X \times Y \to R$ be a continuous function, and $D : X \Rightarrow Y$ be a compact-valued and continuous multifunction. Define m and S as in Theorem 3.

(1) Then m is continuous on X, and S is compact-valued and upper semi-continuous on X.

(2) If $u(x,\cdot)$ is quasi-concave in y for each $x \in X$, and D is convex-valued, then S is convex-valued on X.

(3) If $u(x,\cdot)$ is strictly quasi-concave in y for each $x \in X$, and D is convex-valued on X, then S is continuous function on X.

(4) If u is concave on $X \times Y$, and D has a convex graph, then m is concave function and S is convex-valued multifunction on X.

(5) If u is strictly concave on $X \times Y$, and D has a convex graph, then m is strictly concave function on X and S is continuous function on X.

Recall that the graph of the multifunction D, denoted by $gr(D)$, is the set $gr(D) = \{(x, y) \in X \times Y : y \in D(x)\}$.

Remark 4. It is important to note the following two facts [7, Examples 9.15 and 9.16]:

(1) S is only upper semi-continuous, and not necessarily also lower semi-continuous.

(2) The continuity of u on $X \times Y$ cannot be replaced with one of separate continuity, i.e., that $f(\cdot, y)$ is continuous on X for each fixed $y \in Y$ and that $f(x,\cdot)$ is continuous on Y for each fixed $x \in X$.

Obviously, the following theorem is true.

Theorem 5. Let $X \subset R^n$ and $Y \subset R^m$, $u : X \times Y \to R$ be a continuous function, and $D : X \Rightarrow Y$ be a compact-valued and continuous multifunction. Define m and S as in Theorem 3. If $|S(x)| = 1$ for all $x \in X$, then m and S are two continuous function on X.

Remark 5. Let $X \subset R^n$. A connector for X is a multifunction $C : X \Rightarrow X$ such that $x \in C(x)$ for all $x \in X$; C is called compact connector when $C(x)$ is compact for all $x \in X$.

This remark allows us to introduce the following statement.

Theorem 6. Let $X \subset R^n$, $u : X \to R$ be a continuous function, and D be compact continuous connector for X. Then, the function $m : X \to R$ defined by $m(x) = \max\{u(y) \mid y \in D(x)\}$ is continuous on X, and the multifunction $S : X \Rightarrow X$ defined by $S(x) = \{y \in D(x) \mid u(y) = m(x)\}$ is compact-valued and upper semi-continuous on X.

6. The Indirect Utility Function

The utility function u is defined over the consumption set X of economic agent and it is called the direct utility function. We know that the economic agent (or consumer) choose a utility-maximizing bundle $x(p, m_0)$ for given prices p and income m_0. The level of utility achieved when $x(p, m_0)$ is chosen thus will be the highest level permitted by the budget constraint facing p and m_0, and can be denoted by

$$v(p, m_0) = \max\{u(x) \mid \langle p, x \rangle = m_0\}.$$

The function $v : R_+^m \times R_+ \to R$ that given us the maximum utility achievable at given prices and income is called the indirect utility function. The properties of this function are summarized in the following statement.

Theorem 7 (Properties of the Indirect Utility Function). If the utility function u is continuous and monotonic on R_+^m, $p \in R_+^m$ and $m_0 > 0$, then the indirect utility function v has the following properties:

(1) $v(p, m_0)$ is nonincreasing in p and $v(p, m_0)$ is nonincreasing in m_0.

(2) $v(p, m_0)$ is homogeneous of degree 0 in (p, m_0).

(3) $v(p, m_0)$ is quasi-convex in p.

(4) $v(p, m_0)$ is continuous in $R_{++}^m \times R_+$.

Proof. (1) Let $p, p' \in R_+^m$, $p' \geq p$, $B = \{x \in X \mid \langle p, x \rangle \leq m_0\}$ and $B' = \{x \in X \mid \langle p', x \rangle \leq m_0\}$. For $p' \geq p$ it follows that $B' \subset B$. Hence, the maximum of u over B is at least as big as the maximum of u over B'.

The argument for m_0 is similar.

(2) If prices and income are both multiplied by a positive number, then budget set does not change at all. Thus, we obtain $v(t.p, t.m_0) = t^0.v(p, m_0)$ for every $t > 0$.

(3) Let $p, p' \in R_+^m$ and $t \in (0;1)$. Suppose p and p' are such that $v(p, m_0) \le v_0$, $v(p', m_0) \le v_0$ and $p'' = t.p + (1-t)p'$. Define the budget sets $B = \{x \in X \mid \langle p, x \rangle \le m_0\}$, $B' = \{x \in X \mid \langle p', x \rangle \le m_0\}$ and $B'' = \{x \in X \mid \langle p'', x \rangle \le m_0\}$. We will prove that $B'' \subset B \mathbf{U} B'$. Let assume that is not true. Then we must have $\langle p, x \rangle > m_0$ and $\langle p', x \rangle > m_0$. Multiplying the first inequality by t and the second by $1-t$ and then summing, we find that $\langle p'', x \rangle > m_0$, which contradicts our original assumption. Hence, B'' is contained in $B \mathbf{U} B'$, the maximum of u over B'' is at most as big as the maximum of u over $B \mathbf{U} B'$. As a result we obtain $v(p'', m_0) \le v_0$, i.e. $v(p, m_0)$ is quasi-convex in p.

(4) This follows from the Maximum Theorem, see also Theorems 4 and 5.

The theorem is proven.

References

1. Barten, A., V. Bohm. Consumer theory, Handbook of mathematical economics vol. II, North-Holland Publishing Co, 1982, 381-429.
2. Berge, C. Topological Spaces, Including Treatment of Multi-Valued Function, Vector Spaces, and Convexity, Oliver and Boyd, Edinburgh, 1963.
3. Klein, E., Mathematical methods in theoretical economics, Academic Press, 1973.
4. Slavov, Z. Weak and strong optimality and increasing concave functions, Applied Mathematics and Computation 135 (2003), 517-529.
5. Slavov, Z. Structure of the Pareto optimality set with fixed total resources and consumption sets, Applied Mathematics and Computation 154 (2004), 75-81.
6. Slavov, Z. Scalarization techniques or relationship between a social welfare function and a Pareto optimally concept, Applied Mathematics and Computation 172 (2006), 464-471.
7. Sundaran, R. A First Course in Optimization Theory, Cambridge University Press, 1996.

Zdravko Dimitrov Slavov
Varna Free University, Bulgaria
slavovibz@yahoo.com

ON THE CONCEPTS FOR THE CONTINUOUS SELECTIONS AND APPLICATIONS

Zdravko Slavov, Christina Evans

***Abstract:** In this article we study the role of the continuous selections in mathematical research work. The theory of continuous selections of multi-functions is an intensively developing research area in recent decades, and there have numerous applications in fixed point theory, convex analysis, game theory and other diverse branches of modern mathematics.*

*   **Keywords:** *continuous selection, efficient, multi-objective programming, fixed point*

1. Introduction

The theory of continuous selections of multifunctions is an intensively developing research area in recent decades, and there have numerous applications in fixed point theory, convex analysis, game theory and other diverse branches of modern mathematics. Among a number of continuous selection theorems, Michael's selection theorem is well-known and very basic in many applications [6]. Michael proved that any lower semi-continuous multifunction $F : X \Rightarrow Y$ with closed and convex valued admits a continuous selection, where X is a paracompact topological space and Y is a Banach space.

Let d be the Euclidean metric on R^m and t be the topology induced by d. In a topological space (R^m, t) and by analogy in R^n, for $X \subset R^m$ and $F : X \Rightarrow R^n$ we recall some definitions.

Definition 1. (a) The multifunction F is called to be upper semi-continuous (or usc) at a point $x \in X$ if and only if for each open sets $V \subset R^n$ such that $F(x) \subset V$, these exists a set U of t containing x such that $y \in U \mathbf{I} X$ implies $F(y) \subset V$. F is usc on X if and only if F is usc at each $x \in X$.

(b) The multifunction F is called to be lower semi-continuous (or lsc) at a point $x \in X$ if and only if for each open sets $V \subset R^n$ such that $F(x) \mathbf{I} V \neq \varnothing$, there exists a set a set U of t containing x such that $y \in U \mathbf{I} X$ implies $F(y) \mathbf{I} V \neq \varnothing$. F is lsc on X if and only if F is lsc at each $x \in X$.

(c) The multifunction F is called to be continuous at a point $x \in X$ if and only if F is both usc and lsc at $x \in X$. F is continuous on X if and only if F is continuous at each $x \in X$.

Note that the image of each point by a multifunction is nonempty.

Definition 2. Let $M \subset R^n$ be a nonempty set. A point $x \in M$ is called an efficient point of M if and only if there does not exist a point $y \in M$ such that $y_i \geq x_i$ for all $i \in \{1,2,...,n\}$ and $y_k > x_k$ for some $k \in \{1,2,...,n\}$. The set of the efficient points of M is denoted by $Eff(M)$ and it is called an efficient set.

It is known that $Eff(M)$ is nonempty when M is compact.

Remark 1. Consider a multifunction $F : X \Rightarrow R^n$.

(1) By a selection f of F, we mean a function $f : X \to F(X)$ such that $f(x) \in F(x)$.

(2) Let $F(x)$ be compact for all $x \in X$. By an efficient selection f of F, we mean a function $f : X \to F(X)$ such that $f(x) \in Eff(F(x))$.

2. Continuous selection and multi-objective programming

In this section, we will consider an application of the concept of the continuous selections in optimization theory and multi-objective programming [9].

Theorem 1. Let $X \subset R^m$ and $F : X \Rightarrow R^n$ be a compact-valued and convex-valued continuous multifunction. Then, there exists a continuous efficient selection $f : X \to F(X)$, i.e. $f(x) \in Eff(F(x))$ for all $x \in X$.

Remark 2. We often use the Maximum Theorem (or Berge Maximum Theorem) as a mathematical tool in optimization and nonlinear analysis. In the proof of Theorem 1 we will use this theorem in its classical variant "Let $S \subset R^n$ and $\Theta \subset R^m$, $g : S \times \Theta \to R$ be a continuous function, and $D : \Theta \Rightarrow S$ be a compact-valued and continuous multifunction. Then, the function $g^* : \Theta \to R$ defined by $g*(q) = \max\{g(x,q) \mid x \in D(q)\}$ is continuous on Θ, and the multifunction $D^* : \Theta \Rightarrow S$ defined by $D*(q) = \{x \in D(q) \mid g(x,q) = g*(q)\}$ is compact-valued and upper semi-continuous on Θ" [1] [11, Theorem 9.14].

Proof of Theorem 1. Let $Y = F(X)$ and $x \in X$. For each $i \in \{1,2,...,n\}$ we define a function $m_i : X \to R$ by $m_i(x) = \max\{y_i \mid y(y_1, y_2,..., y_n) \in F(x)\}$. It is easy to see that these functions $\{m_i\}_{i=1}^{n}$ are continuous on X, see Remark 2. As a result we

get a continuous function $m : X \to R^n$ such that $m(x) = (m_1(x), m_2(x), ..., m_n(x))$ for all $x \in X$.

Now, let define a function $g : Y \times X \to R$ by $g(y, x) = -\sqrt{\sum_{i=1}^{n}(y_i - m_i(x))^2}$ for all $y(y_1, y_2, ..., y_n) \in F(x) \subset Y$ and for all $x \in X$, i.e. $g(y, x) = -d(y, m(x))$. It is easy to show that g is continuous on $Y \times X$.

We also see that either $m(x) \notin F(x)$ or $m(x) \in \partial F(x)$.

(1) Let $m(x) \notin F(x)$. From the set $F(x)$ is compact and convex it follows that there exists a unique $z \in F(x)$ such that $g(z, x) = \max\{g(y, x) \mid y \in F(x)\}$ and $g(z, x) < 0$. In this case, we also see that $z \in \textit{Eff}(F(x))$.

(2) Let $m(x) \in \partial F(x)$. This means that $d(m(x), F(x)) = 0$ and $d(m(x), y) > 0$ for all $y \in F(x) \setminus \{m(x)\}$. In this case, let $z = m(x)$. Clearly, $z \in F(x)$ is unique, $z \in \textit{Eff}(F(x))$ and $g(z, x) = \max\{g(y, x) \mid y \in F(x)\}$.

From Remark 2 it follows that $g^* = g(z, \cdot)$ is a continuous function on X and the multifunction $F^* : X \Rightarrow Y$ defined by $F^*(x) = \{y \in F(x) \mid g(y, x) = g(z, x)\}$ is compact-valued and upper semi-continuous on X.

Note that $|F^*(x)| = 1$ and $z \in F^*(x)$. This allow us to define a function $f : X \to R^n$ by $z \in f(x)$, i.e. $f = F^*$. As a result we deduce that f is continuous function on X, $f(x) \in F(x)$ and $f(x) \in \textit{Eff}(F(x))$ for all $x \in X$.

The theorem is proven.

3. Continuous selection and fixed point properties

In this section, we will consider an application of the concept of the continuous selections in fixed point theory.

Definition 3. The topological space Y is said to have the fixed point property if and only if every continuous function $h : Y \to Y$ from this set into itself has a fixed point, i.e. there is a point $x \in Y$ such that $x = h(x)$.

Remark 3. From a more formal viewpoint, a retraction is a function $r : X \to Y$ such that $r \mathbf{o} r(x) = r(x)$ for all $x \in X$, since this equation

says exactly that r is the identity on its image. Retractions are the topological analogs of projection operators in other parts of mathematics. Clearly, every deformation retract is a retract, but in generally the converse does not hold [5] [12].

Let us consider a multifunction $j : Y \Rightarrow Y$. Let it be upper semi-continuous with a nonempty, compact and convex image, shortly we say that j is *cusco*.

Definition 4. The topological space Y is said to have the Kakutani fixed point property if and only if every *cusco* $j : Y \Rightarrow Y$ has a fixed point, i.e. there is a point $x \in Y$ such that $x \in j(x)$.

Lemma 1 [11, Theorem 9.31 - Schauder's Fixed Point Theorem]. Let $h : S \to S$ be a continuous function from a nonempty, compact and convex set $S \subset R^n$ into itself, then h has a fixed point.

Lemma 2 [11, Theorem 9.26 - Kakutani's Fixed Point Theorem]. Let $S \subset R^n$ be a nonempty, compact and convex set and the multifunction $j : S \Rightarrow S$ be *cusco*, then j has a fixed point.

Remark 4. A property is called a topological property if and only if an arbitrary topological space X has this property, then Y has this property too, where Y is homeomorphic to X. Compactness, contractibility and the fixed point properties (the fixed point property and the Kakutani fixed point property) are topological properties.

Remark 5. The fixed point and the Kakutani fixed point properties of sets are preserved under retraction [5] [8] [10]. This means that the following statements are true: if set X has the fixed point property and Y is a retract of X, then set Y has the fixed point property too; if set X has the Kakutani fixed point property and Y is a retract of X, then set Y has the Kakutani fixed point property too.

Remark 6. The Kakutani fixed point property is very closely related to the fixed point property. If $S \subset R^n$ has the Kakutani fixed point property, then since any continuous function from S into itself can be viewed as a *cusco* it follows that set S will also have the fixed point property. For more information see also Lemmas 1 and 2.

Remark 7. Let $S \subset R^n$ be compact. It can be shown that set S having the Kakutani fixed point property is equivalent to S having the fixed point property. Remark 6 has shown that if S has the Kakutani fixed point property, then S has the fixed point property. Now, let $j : S \Rightarrow S$ be *cusco*, S have the fixed point property and $gph(j) = \{(x, y) \in S \times S \mid y \in j(x)\}$. From Cellina's Theorem it follows

that there is an approximate continuous selection h of j [2] [3]. That is, for each $k \in N$ there exists a continuous function $h_k : S \to S$ such that $d((x, h_k(x)), gph(j)) < \dfrac{1}{k}$ for all $x \in S$. From the assumption that S has the fixed point property, it follows that each function h_k has a fixed point $x_k \in S$. As a result we get a sequence $\{x_k\}_{k=1}^{\infty} \subset S$ such that $d((x_k, x_k), gph(j)) < \dfrac{1}{k}$, i.e. the point (x_k, x_k) approaches the set $gph(j)$. The set S is compact, implying that there exists a convergent subsequence $\{x'_{m(k)}\}_{m(k)=1}^{\infty} \subset \{x_k\}_{k=1}^{\infty}$ such that $\lim\limits_{k \to \infty} x'_{m(k)} = x_0 \in S$. We also see that $d((x'_{m(k)}, x'_{m(k)}), gph(j)) < \dfrac{1}{m(k)}$. But j is *cusco*, then $gph(j)$ is closed. Taking the limit as $k \to \infty$ we have $m(k) \to \infty$ and obtain $\lim\limits_{k \to \infty}(x'_{m(k)}, x'_{m(k)}) = (x_0, x_0) \in gph(j)$. This means that $x_0 \in S$ is a fixed point for j, see also [4] and [10]. Finally, we find that S has the Kakutani fixed point property.

Remark 8. As we have shown in Lemma 1, if an arbitrary set is nonempty, compact and convex, then it has the fixed point property. In general, it is not difficult to verify that the efficient set is nonconvex, but it is compact and contractible [7] [10]. Thus among nonconvex sets, compactness and contractibility do not have direct relationship with the fixed point property. There are examples of compact and contractible sets which do not have the fixed point property. It is not known what types on nonconvex sets have this property.

References
1. C. Berge. Topological Spaces, Including Treatment of Multi-Valued Function, Vector Spaces, and Convexity, Oliver and Boyd, Edinburgh, 1963.
2. J. Borwein, A. Lewis. Convex Analysis and Nonlinear Optimization: Theory and Examples, Springer, 2000.
3. A. Cellina. Fixed Point of Noncontinuous Mapping, Atti Accad. Naz. Lincei Rend. 49 (1970), 30-33.
4. J. Franklin. Methods of Mathematical Economics: Linear and Nonlinear Programming, Fixed Point Theorems, Springer, 1980.
5. A. Hatcher. Algebraic Topology, Cambridge University Press, 2002.
6. E. Michael. Continuous selections, I. Ann. Math. (63) 1956, 361-381.

7. Z. Slavov. The Fixed Point Property in Convex Multi-Objective Optimization Problem, ACTA UNIVERSITATIS APULENSIS (15) 2008, 405-414.
8. Z. Slavov, C. Evans. Fixed Point Properties of the Pareto Sets in Multi-criteria Maximization, International Conference ISK 2010, Varna 24 – 26 June 2010, 150-158.
9. Z. Slavov. Fixed Point Properties in Continuous Engineering Maximization, International Conference RaDMI 2010, Donji Milanovac 16 - 19 September 2010, Serbia, 1224-1228.
10. Z. Slavov, C. Evans. Compactness, Contractibility and Fixed Point Properties of the Pareto Sets in Multi-Objective Programming, Applied Mathematics (2) 2011, 556-561.
11. R. Sundaran. A First Course in Optimization Theory, Cambridge University Press, 1996.
12. A. Wilansky. Topology for Analysis, Dover Publications Inc., 1998.

Zdravko Dimitrov Slavov
Varna Free University, Bulgaria
slavovibz@yahoo.com

Christina Slavova Evans
The George Washington University, USA
mrsevans@yahoo.com