REVISTA ECONOMICĂ

Supplement No. 5/2012

Journal of economic-financial theory and practice

CNCSIS reference: B+, Code 478
Frequency: 6 issues/year

Indexed in:
RePeC
Ulrich`s Periodicals Directory
DOAJ - Directory of Open Access Journals

Lucian Blaga University of Sibiu
No.10, Victoriei Bd.
Postal Code 550024, Sibiu, Romania

Center for Economic Research
Lucian Blaga University of Sibiu
No.17, Calea Dumbrăvii Street
Postal Code 550324, Sibiu, Romania
Phone/Fax: +40 (269) 23.58.79

G.P.E. «Continent»
No.2, Dr. Ion Rațiu Street
Postal code 550012, Sibiu, România
Phone: 40-269-243 210

Mail: revista.economica@ulbsibiu.ro
Web: http://economice.ulbsibiu.ro/revista.economica

ISSN: 1582-6260

THE IMPACT OF SEMANTIC WEB TO BUSINESS INTELLIGENCE

BERTA Dora-Anca
Ph.D. student, Faculty of Economics and Business Administration /Department of Accounting, Statistics and Information Systems, "Al. I. Cuza" University, Iasi, Romania, anca.berta@feaa.uaic.ro

Abstract: *Business Intelligence are high-level decisions support applications, their role being to provide to management, based on information from the company, a picture of its situation at some time but also on the medium and long term, bring forward taking competent decisions. The purpose of this research is to identify, describe and analyse if there is the impact on the Semantic Web solutions for Business Intelligence and how it manifests. The appearance of a new generation applications, called Semantic Business Intelligence will confirm that the Semantic Web has made its mark on them.*

Keywords: business intelligence, semantic web, semantic technologies, data integration

JEL classification: L86, O33

1. Introduction

Nowadays, amount of data from organizations grows exponentially. More information means more competition. In a period of information explosion, managers, users, and not only have to take optimal decisions in a short time as possible. Competitive organizations are permanently prepared to meet the challenges of other actors from the business environment.

In the economic field the ancient dictum “All the roads leads to Rome” implies the fact that the majority of analysts from this area tries to identify the cutting-edge technologies which will change the future in business by using the concept of Business Intelligence.

Although the concept of Business Intelligence appears into IBM Journal in 1958 as "automated system used to disseminate information in various sectors of organizations" (Berta, 2012, p.62), in the scientific literature began to be circulated only since 1989 when analyst Howard Dresner, Garter Group company employee defines it as it "an umbrella term that means a set of concepts and methods to improve the process of making business decisions using decision assistance systems based on facts" (Power , DJ, 2010).

Business Intelligence is not only a concept, but also a valuable instrument in the decision-making management. In this way, B.I. will allow the processing, the analysis and the forecast of data from an organization. Companies use B.I. for its use in making a basis for the process of decision-making.

According to market surveys conducted by Gartner Group, at present, revenues from the sale of these solutions exceeding the fabulous sum of $ 12 million (Pettey, C., van der Meulen, R., 2012), the world's leading BI market are SAP, Oracle, Microsoft, IBM SAS Microstrategy, Actuate, Panorama Software, TIBCO Software, etc.

This research aims to explain a new trend observed in the structure of BI solutions. The approach is descriptive-explanatory, proposing to bring some clarifications regarding the “semantic" side of Business Intelligence solutions. The importance of this approach is justified because we have in mind the implication of the internet users, consciously or not, in the phenomenon of semantic integration.

2. Research methodology

The issue proposed to be analysed in this research occurs as a consequence of scientific observation and analysis of the semantic web impact on data and, by default, on the applications which produce, mange, process or analyse these data. B.I. is the most-known instrument in the review, the reprocessing and the design of data. Thus, the study of the relation between the semantic web and B.I. is important.

The **hypothesis** in this analysis is “*There is a significant impact of the semantic web on B.I. solutions”*.

In the construction of this hypothesis we used the *concordance method*. This method explains the effects of a phenomenon causes according to various factors highlighted (Zaiţ D., şi Spalanzani, A., 2006, pp. 145-146). On a concrete level, in this case the marked phenomenon was the development of the

semantic web. The evolution of semantic the web was achieved through semantic integration of data which determined a rethink of B.I. solutions in such a manner that it can analyze, process and synthesize, in addition to traditional data and semantic data integrated which have a different form and structure.

The inferring type is *causally* because explain causes but mainly effects of crystallization and the Semantic Web development and also *deductive* whereas in order to draw conclusions will be realized a synthesis of the analysis. Hence this research will have a deductive scientific undertaking, proved by the fact that starting with a true statement proposition will get another true statement proposition. The approach used is qualitative, comparative functionalist type (Wacheux, 1996 *apud* Zaiț D., și Spalanzani, A., 2006, pp. 176-177) whereas analyzes the deterministic influence of a situational context, in this case the introduction of new technologies in certain applications. The appearance of a new generation of

Business Intelligence generically called Semantic Business Intelligence based on on the Semantic Web specific technologies will confirm the proposed hypothesis.

Data was collected using specialized databases (ScienceDirect, Emerald Database, Springerlink), websites that are relevant to the topic (e.g. http://semanticweb.com, http://wi-consortium.org/), scientific magazines (e.g. Journals of Information Technology and Systems, Information System Management, etc.) and other specific research.

Materials selection we attend to the following aspects: the true (trust) source of originating, the magazine's prestige where it was published the material (ISI articles and BDI items), their actuality and relevance to the treated topic.

3. Web and Business Intelligence – evolutions and intersections

According to Berners-Lee, the Web is defined as a distribution of multimedia information both locally and globally (Berners-Lee, T., 1999). Web and HTML appears in 1991, when Tim Berners-Lee creates first web page. He said about web "that is more a social creation than a technical one" (Anghel, T., 2009, p 83, cited in Berners-Lee, 1999) fact confirmed through the social networks and applications widely used by users around world.

In terms of BI, Web 1.0 has not influenced in any way the existence and/or develops itself. Independently, it crossed several stages, moving from Operational BI to Real-Time BI.

Since 2004, in the literature has appeared a new concept, Web 2.0 or Social Web, where Internet users were not only consumers but also producers of online information of data through blogs, wikis, podcasts, etc. (O'Reilly, 2005). The extent that took the use of web 2.0 and social networks has left its mark on BI solutions, causing the implementation of social modules that include messages and posts on social networks in order to analyze their influence on companies (Berta, D.-A., 2011, p. 3). In this stage appeared applications like Social CRM and Social Business Intelligence.

In 2001, into Scientific American magazine, Tim Berners-Lee, James Hendler and Lassila Time delineated a vision under crystallization related to into Semantic Web; they saw it as a highly interconnected network of data that could be easily accessed, understood and used by any desktop or portable computer (Berners et al, 2001).

4. Semantic technologies used in data integration

In the vision of Berners-Lee, Web 3.0 is an extension of the current web, the difference occurring in the fact that the information would get a definite meaning and a purpose for computers that process. For this to happen, information should be expressed in a specific way to be seen properly and the computer.

Technologies which can make Berners-Lee's vision to become reality must be based on a common language for data representation, in order to be understood by all types of software agents, through ontologies to interpret in terms of common information bases disparate data and rules allowing software agents to reason the information described in these terms. The most known and used semantic technologies are used RDF, RDF Schema, OWL, SKOS, SPARQL, RIF and GRDDL.

Given the fact that will be mentioned in subsequent chapters, is necessary to describe two of technology-namely RDF and OWL.

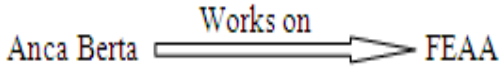
For a computer, to know what to do base on Web content that interacts with, it would have to "catch" the meaning context, this thing happens using the technologies mentioned above.

RFD (Resource Description Framework) - signifies a framework for data representation on the web, intranet extranet. Derived from XML standard uses URI (Uniform Resource Identifier) to identify a resource and to make statements on that resource (Klyne, G., Carroll, J., 2004). The way to make statements about web resources by RFD takes the form of a triplet consists of subject, predicate and object, where (Buraga, SC, 2009):

- ✓ The subject represents the resource (identified by a URI)
- ✓ The object represents values of the resource
- ✓ The predicate – specify the nature of relationship between subject and object.

By creating these triplets, containing subject, predicate, and object, RDF allows computers to form logical sentences based on a combination of subjects and objects. An example of a triplet could be:

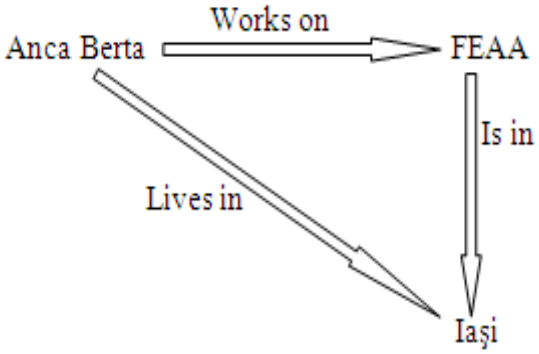
Figure no. 1 Example of triplet



where "Anca Berta" is subject, the predicate is "working to" and "FEAA" is subject.

Each triplet RDF makes a graph arch and several triples combined form a graph.

Figure no. 2 Example of a graph



As observed, should be noted that the RDF is limited to provide a model and syntax for describing resources, but this is not semantics (meaning) resources, the real semantic being obtained by RDFS (RDF Schema) OWL (Web Ontology Language).

Also, RDF statements aimed instances (individuals) but according to requirementsis required vocabularies that can be include instances, specifically by defining a taxonomy of resources, including classes, sub-classes super classes that can be realized by RDF Schema (Brickley Guha, 2004). RDF Schema provides sense to RDF resources by adding semantics of predicates specifying properties, values and use restrictions that may have a certain time. So, RFD scheme is a framework that allows specification of a vocabulary of terms and relations between them, in a given field (Buraga, SC, 2009).

OWL (Web Ontology Language) - according to Gruber, an ontology is a formal explicit specification of a shared conceptualization (Gruber, T., 2009). Also, OWL, is a way of representing knowledge as a common vocabulary that can be used to model a domain, consists of sets of objects / of existing concepts and properties and their interrelations (Arvidsson, F. Flycht-Eriksson, A., 2002). The elements used have the same meaning for all those who use them.

In terms of computing, ontology is a conceptualization of domain knowledge in a format intended to be processed by computer, format the model entities, attributes, relationships and axioms.

In recent years, the trend was to adopt ontologies as in many scientific communities as a way to share, reuse and process different fields of knowledge. Ontologies are now centers of many applications such as scientific knowledge portals, information management, electronic commerce and semantic web.

There are lots of plug-ins to be imported ontologies in different formats into a specific software called Protege including DAG-EDIT, XML, RDF, and OWL.

5. Sematic Web - premise of Semantic Business Intelligence appearance

To discuss about semantic integration, must first be clarified notion of "semantic". According to specialists (Ouksel Sheth, 1999, Fikes McGuinness, 2001; Uschold, 2002) semantic typology applicable to computer (Uschold, 2002, pp. 3-4) take the following forms: real-word semantics, axiomatic semantics and model-theoretic semantics.

Real world semantics (real world) refers to "mapping objects computational models of real world languages and the various situations that involve human interpretation, understanding use of data or information" (Ouksel Sheth, 1999, p 4). Specifically, this case is about semantic item level (item) that could be a label, term, phrase or set of expressions to represent a range of real world using real-world concepts.

Axiomatic semantics applied to a specific language (Fikes McGuinness, 2001) "the mapping of the set of descriptions in [That] language Expressed in the logical theory into first-order predicate calculus. ". Axiomatic semantics is found in the Resource Description Framework (RDF), RDF Schema (RDF-S) Damla + OIL.

Model-theoretic semantics (model-theoretically Semantics) of a language that language assumes that addresses "the world" describes the minimum requirements that a "world" must meet to assign an appropriate by of each phrase of the language used (W3C, 2002). Is used as a technical tool to determine the operations proposed language retains its meaning; characterize valid conclusions to be drawn from a given set of expressions, independent of the symbols.

Diversified data format is one of the reasons that led to a new generation of BI because BI solutions currently used applies only to structured data that are significantly less than unstructured data - according to Paquet unstructured data represents about 80% of all existing data worldwide (Paquet, R., 2010).

Applications currently used are or should be designed so as to handle heterogeneous data as the source format, in the literature data (the economic) are divided into three categories (Wang, R., 2010):

a) structured data: data is stored using traditional DBMS of applications like ERP, CRM, SCM;
b) semistructured data: is the data collected from RSS feeds, spreadsheets and XML documents;
c) unstructured data: derived from emails, blogs, social networks, mobile phones, etc. and can take various forms including: audio, video or images.

	Structured	**Semi-structured**	**Unstructured**
Explanation	Describes data structure (the type, the schema)	Describes data effectively	Present data without any description
Data syntax	<person> Anca Berta </person> <simpleType name= “person”> <restriction base=”string”> <pattern value=”\w{20}”/> </restriction> </simpleType>	<person>Anca Berta</person>	Anca Berta

For clarification, I will take an example to see the differences between a structured, a semi-structured and an unstructured data.

Table nr. 1 Example of different types of data

Processed after example of: Kamiski, R., Panas, T., Semantic Web, 2002

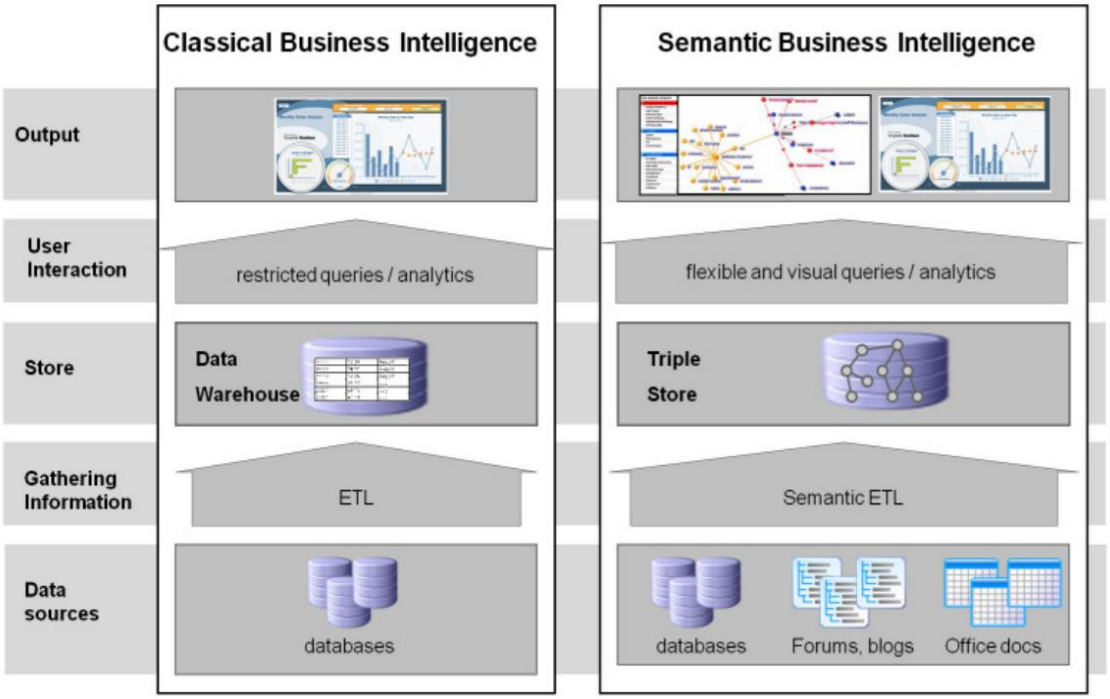
Semantic data integration is performed using specific technologies and must be seen through the benefits it brings:

- ✓ data is converted into a common format
- ✓ disparate data are grouped together in a place by sense fields, etc.
- ✓ integrated data mining and search within them is interactive and easy

6. Semantic Business Intelligence: a new generation of BI

For years where minor changes were made in the design and use of Business Intelligence's site, nowdays we are witnessing a major change of BI solutions. Currently, we hear more and more about the appearance of a new generation of BI solutions, called Semantic BI. SBI is not intended to replace traditional BI, but is intended to be an extension of it, the difference being much richer data format to be processed.

Figure nr 3 Traditional vs. Semantic Business Intelligence



Source: http://www.cubist-project.eu/uploads/media/CUBIST_factsheet_extended.pdf

Looking at figure no 4 we observe that the differences between the two dotted generation Business Intelligence specialists:

- ✓ data source
- ✓ data storage
- ✓ data analysis

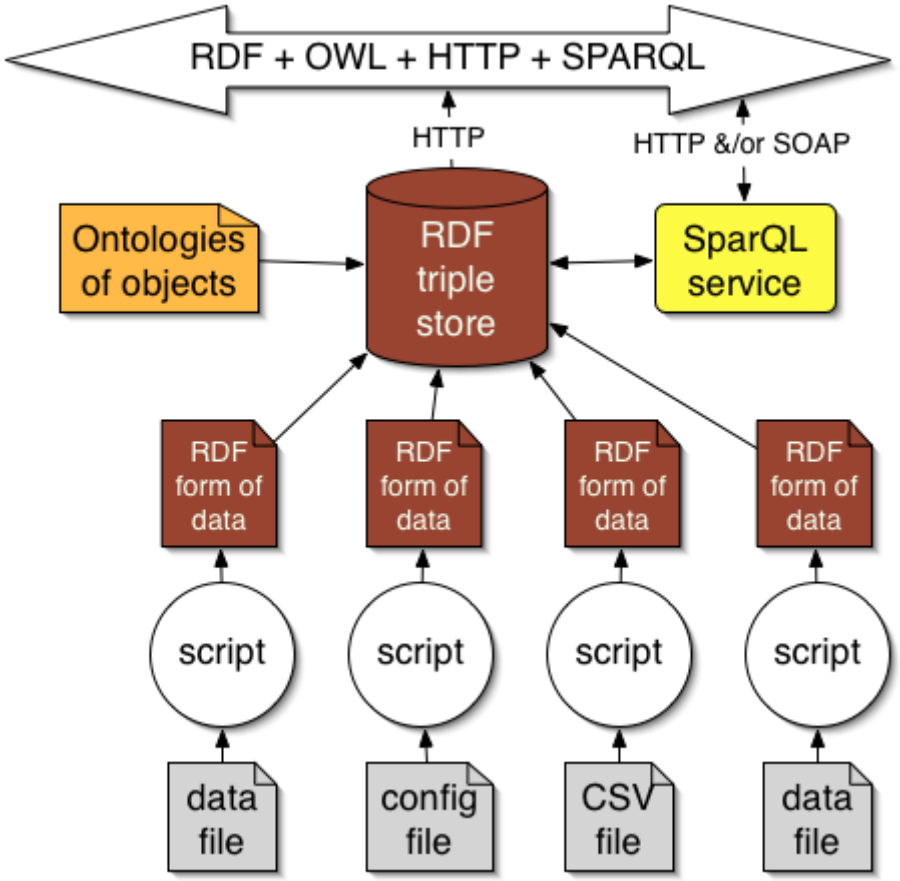
Data source: semantic data integration advantage comes from the abundance and variety of data can be used, the Internet currently prevailing unstructured data, data that were not counted so far lacking to exploit appropriate technologies. Using views from blogs, RSS data feeds, and XML sites, news sites, a company can know exactly its market position relative to competitors or in relation to the clients or business partners.

Data storage: in traditional BI, data is stored in a Data Warehouses. A data warehouse in Inmon's concept is a collection of integrated databases, subject-oriented and designed to provide information necessary for decision making (Airinei D., 2003, p. 3). For Semantic BI, the data is stored in triple stores.

A triple store is a framework used to store query RDF data. It provides a persistent storage mechanism access to RDF graphs. The first triple stores appeared in early 2000 and it was Jena Sesame. Today, on the market there are many such applications, the most famous being Sesame, Mulgara, Garlik JXT, 3Store, YARS2, BigOWLIM, Virtuoso, AllegroGraph, BigData, TDB Jena, Jena SDB, Kowari, RDFGateway.

Eventually we see triple stores as baye to give some specially designed to store triplets, where a triplet is a data defined by three cooordonate: subject, predicate and object. In figure no 4, we can a possible structure of a triple store:

Figure no. 4 A triple store schema



Source: Berners-Lee, T., *Semantic Web Concept*, http://www.w3.org/DesignIssues/diagrams/sw-store.png

Data analysis: if the traditional BI analysis is used for data mining., In datle SBI will be analyzed using FCA (Formal Concept Analysis). FCA is a method used for data analysis, knowledge representation and information management. Data units analyzed are structured as formal abstractions of concepts of human thought, allowing a comprehensive and easily understandable interpretation (Priss, U., 2007).

7. Conclusion

Business Intelligence provides to organization support to optimally manage information on their so as organization to able to adopt the best decisions for new competitive advantages. In the same time, BI provides users vital information about techniques used in business and thinking of the competitors, their culture, their intentions and ability to implement them.

Research results show that between Semantic Web Business Intelligence is a link due to specific semantic web technologies like triple stores and Formal Analysis Concept based on RDF, RDF Schema OWL and SPARQL, used by new generation of BI applications, which confirms the hypothesis.

Business Intelligence's alignment to this new trend, driven by web development will be neither simple nor accessible. It can discuss even a rethinking of the entire architecture of BI solutions in the integration of both internal and external data to support collaboration. The difficulty comes from the requirement to integrate different data types (including data from social networks), by biggest volume and varied and scattered sources in a simpler and intuitive form to be understand and use.

8. Acknowledgement

This work was supported by the European Social Fund in Romania, under the responsibility of the Managing Authority for the Sectorial Operational Program for Human Resources Development 2007-2013 [grant POSDRU/CPP 107/DMI 1.5/S/78342].

9. References

- Anghel, T., (2009), *Instrumente Web 2.0 pentru profesori*, Editura All, Bucureşti;
- Airinei, D., Depozite de date, Editura Polirom, Iaşi, 2003;
- Berners-Lee, T., (1999), *Weaving the Web*, Orion, Cambridge, available at http://www.w3.org/People/Berners-Lee/Weaving/;
- Berners-Lee, T., (1998), *Semantic Web Road map*, available at http://www.w3.org/DesignIssues/Semantic.html;
- Berners-Lee, T., Hendler, J., Lassila O., (2001), *The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, Revista Scientific American, available at http://www.scientificamerican.com/article.cfm?id=the-semantic-web;
- Berta, D.-A., (2012), *Business Intelligence in education*, Leveraging Technology for Learning, Volume II, Proceedings of The 8th International Conference eLearning and Software for Education, Bucharest, 26-27 April, pp. 62-67;
- Berta, D.-A., (2011), *From Marketing Viral to Social CRM*, Analele Universitatii Ovidius, Seria: Stiinte economice, Nr. Vol. XI, nr. 2/2011, Editura , 2011, ISBN ISSN 1582-9383, pp. 87-92;
- Buraga, S. C., (2009) *RDFSchema and Skos*, Suport de curs, available at http://profs.info.uaic.ro/~busaco/teach/courses/wade/presentations/web05RDFSchema-SKOS.pdf
- Hudson, D., (2011), *The Semantic Social Business Phase of Enterprise 2.0*, available at http://webtechman.com/blog/2011/05/16/enterprise-3-0-semantic-social-business/;
- Lachev, T., (2010), *Business Intelligence Semantic Model – The Good*, The Bad, and the Ugly, available at http://prologika.com/CS/blogs/blog/archive/2010/11/13/business-intelligence-semantic-model-the-good-the-bad-and-the-ugly.aspx;
- Mattison, S., (2010), *Microsoft Business Intelligence Conference at TECH-ED North America*, New Orleans, Louisiana, available at http://blogs.adatis.co.uk/blogs/simon/default.aspx;
- Mazoue, D., , Imm, E., Astito, A., (2010*), Questions and Answers: SAP BusinessObjects* BI 4.0 and the New Semantic Layer for OLAP, available at http://www.sdn.sap.com/irj/scn/go/portal/prtroot/docs/library/uuid/b08143a1-eea3-2e10-da94-f6c5b122e0f4?QuickLink=index&overridelayout=true&51887500377508 ;
- Medin, D., Schwartz, H., Blok, S., Birnbaum, L., (1999), *The semantic side od decision making*, Psychonomic Bulletin and Review, Nr 6 (4), 562-569;
- Napoli, M., (2008), Ghisi, F. B., Pacheco, R., Todesco, J. L., SBI: *A Semantic Framework to Support Business Intelligence*, OBI'08, Octombrie, Karlsruhe, Germania, ACM 978-1-60558-219-1/10/08
- Oliver, M., (2009), *The Semantic Enterprise*, available at http://www.zdnet.com/blog/collaboration/the-semantic-enterprise/183 ;
- O'Rilley, T., (2005), What Is Web 2.0 - *Design Patterns and Business Models for the Next Generation of Software*, http://oreilly.com/web2/archive/what-is-web-20.html
- Paquet, R., (2010), *Technology Trends You Can't Afford to Ignore*, available at http://www.gartner.com/it/content/1503500/1503515/january_19_tech_trends_you_cant_

afford_to_ignore_rpaquet.pdf ;

- Passant, A., Bojars, U., Breslin, J., Decker, S., (2009), Semantic Enterprise 2.0 - Enabling Semantic Web Technologies in Enterprise 2.0 Environment, Semantic Technologies Conference, San Jose, SUA, available at http://semantic-conference.com/session/1783/ ;
- Patrichi, M., (2009), Microsoft's Sharepoint: *From Zero To Billion In Less Than A Decade*, Proceedings to "Institutions and economic performance" International Conferrence, 20-21 November, Iasi;
- Pettey, C., van der Meulen, (2012), *Gartner Says Worldwide Business Intelligence*, Analytics and Performance Management Software Market Surpassed the $12 Billion Mark in 2011, available at http://www.gartner.com/it/page.jsp?id=1971516;
- Piasevoli, T., (2011), *Business Intelligence Semantic Model*, Microsoft Local WinDays 11 Conference, Croatia-Rovinj, available at http://www.mswindays.com/
- Power, D.J., (2007), *A Brief History of Decision Support Systems*, version 4.0, DSSResources.COM, World Wide Web, available at: http://DSSResources.COM/history/dsshistory.html
- Priss, U.,(2007), *Introduction to and Overview of Formal Concept Analysis*, Annual Review of Information Science and Technology, Vol. 40;
- Shannon, V., (2006), *A 'more revolutionary' Web: Interview with Tim Berners-Lee*, International Herald Tribune, available at http://www.iht.com/articles/2006/05/23/business/web.php ;
- Sheth, A., (2002), *Semantic Enterprise – courses*, LSDIS Lab, Computer Science, University of Georgia, available at http://lsdis.cs.uga.edu/S0emanticEnterprise/ ;
- Uschold, M., (2002), *Where are the Semantics in the Semantic Web?*, University of Georgia, available at http://lsdis.cs.uga.edu/SemWebCourse_files/WhereAreSemantics-AI-Mag-FinalSubmittedVersion2.pdf ;
- Zaiţ, D., Spalanzani, A., (2006), *Cercetarea în economie şi management – repere epistemologice şi metodologice*, Bucureşti: Editura Economică;
- ***, CUBIST 2010 Annual Report, available at http://www.cubist-project.eu/uploads/media/CUBIST_257403__annual_report_2010.pdf ;
- ***, http://www.bioontology.org/wiki/images/6/6a/Triple_Stores.pdf

AN EMPIRICAL EXPLORATION IN ASSESSING THE DIGITAL SKILLS FOR ROMANIA

BORISOV Daniela
Professor, Faculty of Management/Department Management, Academy of Economic Studies, Bucharest, Romania, daniela.borisov@yahoo.com

FRĂŢILĂ Laurenţiu
Associate Professor, Faculty of Accounting and Management Information Systems/Department of Accounting and Management Information Systems, Academy of Economic Studies, Bucharest, Romania, laurentiu_f@yahoo.com

TANŢĂU Adrian
Professor, Faculty of Business Administration/Department UNESCO, Academy of Economic Studies, Bucharest, Romania, ad_tantau@yahoo.com

Abstract: *The idea of the paper was inspired by link between the process of digital skills' acquiring and the evolving digital divide. There are referred some evaluations on the Information and Communication Technology (ICT) as related to different facets of computer and internet skills. Some methods are involved to make a picture of the current situation on Romania in this domain: comparative statistical analysis, and the decile approach and z-scores. The countries of interest were Sweden and Finland as very competitive EU state members in the development of the information and knowledge society and economy and the newest EU state members Romania and Bulgaria.*

Key words: digital/e-skills, metrics, divide

JEL classification: O30

1. Introduction

The present paper deals with dynamics and regional distribution of digital skills - a subject of crucial importance in designing the future condition of the collection of competences necessary for the professionals in the current marketplace; it also relates to improving access of any individual in the civil life, including public services such as e-government, e-health.

The link between the digital skill acquiring and the evolving digital divide take forma of a social process that may be built on social inequalities and may lead to the marginalization of individuals and groups as regards shared societal goals.

E-Skills are declared by various forums of the European Commission as one of the top priorities of the information society. It is an obvious trend that all individual, in their professional capacity or simple citizens need to develop and enhance digital skills to be able to participate fully on the labour market and in society. The increasing use of computers in the work place has led to computer literacy being a necessity in a large majority of professions. This calls attention on the possible disequilibria in the mechanisms of demand and supply in respects to digital skills and need appropriate assessment on their developments.

The "secondary" digital divide is seen as stemming from the quality of use. It now "describes the difference between users. There is emerging evidence of a second digital divide: some users lack the skills to take advantage of the more advanced services available on the Internet and are unable to carry out online transactions. From a policy perspective, this secondary divide is particularly significant in terms of online public services such as e-Government and e-Health.

2. Methodological aspects

In order for an individual to become digitally competent, one first has to have access to the ICT. Therefore, access to digital technologies is a crucial environmental factor contributing to digital competence. In the past, this access meant having access to a computer. The basis of digital competence is the acquisition of basic operational ICT skills. Basic operational skills refer to the basic skills needed in order to allow the functional use of ICT. Those skills therefore should include basic operational skills for

computers and the internet. Beyond the ability to perform basic operational ICT tasks, digital competence involves the use of ICT for a multitude of aspects of everyday life.

The working definition of digital literacy adopted in the paper is the skills required to achieve digital competence. Digital literacy is the skills required to achieve digital competence, the confident and critical use of ICT for work, leisure, learning and communication.

Digital (computer) skills are defined as having performed at least one of the following computer-related activities: copying or moving a file or folder, using copy and paste tools to duplicate or move information within a document, using basic arithmetic formulas in a spreadsheet, compressing (or zipping) files, connecting and installing new devices, writing a computer programme using a specialized programming language. Low skills refers to being able to do one or two of these computer-related activities while medium skills requires three or four, and high skills five or all of them (Europe's Digital Competitiveness Report, 2010).

Digital skills are proxied by the narrower category of computer skills; the correlation between internet and computer skills (aggregating all skill levels) exceeds 0.99 (Europe's Digital Competitiveness Report, 2010).

The terms digital divide and e-Inclusion implicitly refer to a dichotomy. With regard to a divide, there are people on either side of it, with regard to inclusion, there are those included as opposed to those excluded. Observers, nevertheless, often make the statement that a specific disadvantaged group has caught up or that another one is falling apart. Such statements often are based on vague or unspecific notions of the dynamic process. For the above mentioned characteristics of diffusion, various measures of inequality have been proposed. These include:

- Percentage point differences – involve the absolute distance between the penetration rates. It is questionable because fails to derive a genuine conclusive statement about the evolving that divide as long as the percentage point difference increases and divides are decreasing once the percentage difference decreases is reasonable.
- Percentage ratios - are the base of the Digital Divide Index (DIDIX) approach which takes on a value of 1 (or 100%) in case of equality in the take up between an at-risk group and the population total. These properties make percentage ratios one suitable metric of choice for e-Inclusion measurement.
- Time Distance - Sicherl (2003) has proposed a metric, the S-distance, which measures the time distance between different diffusion curves. The time distance is to measure the time lag that a unit lagging behind needs to reach the level of an advanced unit.
- Growth rates – proposed by the 2002 US report "A Nation Online" concluding that the (American) digital divide is supposed to be closing because the growth rate in internet access of low income groups was higher than the one of high income groups.
- Odds growth rates - Inspired by Martin's work (2003) arguing that the use of odds ratios is a better indicator to judge the dynamics of a digital divide than growth rates.

Reports relating to the internet and computer usage. A useful tool to monitor the development of population acquiring digital skills, not only in the European space but all over the world is the ICT Development Index (IDI) proposed by the International Telecommunication Union (ITU). The ITU annual report "*Measuring the Information Society 2011*" compute a composite index covering of 11 indicators for ICT access, use and skills, which was designed to measure the level and evolution of ICT developments. The IDI aims to capture the evolution of the information society as it goes through its different stages of development, taking into consideration technology convergence and the emergence of new technologies. The IDI is divided into the following three sub-indices:

• *Access:* captures ICT readiness and includes five infrastructure and access indicators (fixed telephony, mobile telephony, international Internet bandwidth, households with computers, and households with Internet).

• *Use:* captures ICT intensity and includes three ICT intensity and usage indicators (Internet users, fixed broadband, and mobile broadband).

• *Skills:* captures ICT capability or skills as indispensable input indicators. The skills sub-index therefore has less weight in the computation of the IDI compared to the other two sub-indices. In the absence of reliable and comparable data on ICT skills for a large number of countries, the three indicators included in the skills sub-index (adult literacy, gross secondary and tertiary enrolment) are proxies. This sub-index should be considered as an enabler for effective ICT use, but economies that rank highly in this sub-index do not necessarily rank highly in the other sub-indices or the overall IDI.

Table 1. IDI skills sub-index, 2010 and 2008

Country	Rank in 2010 (out of 152)	Score in 2010	Rank in 2008	Score in 2008
Finland	1	9.89	2	9.77
Sweden	19	8.89	14	8.99
Romania	30	8.58	31	8.48
Bulgaria	46	8.09	48	7.92

Source: Measuring the Information Society 2011

3. The empirical analysis. The outlook on Internet access for Romanian citizens

Recently, there are more signs of Romania poor positioning in respect to secondary digital divide. The divide is following the levels of skills and empowerment of users. Skills represent the abilities to use internet services: navigation skills and problem solving skills. Empowerment refers to the effectiveness of use with regard to social relationships and social capital. In this phase, the divide will exist along the variables: education, income, gender, age and period of use (Benchmarking in a Policy Perspective, 2006).

Romania is still lagging behind in the implementation of the Information Society. There is strong correlation between the progress made in Internet regular use and computer and Internet skills. A detailed view on the ICT development is provided by the latest ITU report which gives a description on main ICT indicators reported by ITU.

Table 2. The IT profile for Romania (2008-2010)

Access indicators	**Main (fixed) telephone lines per 100 inhab.**		**Mobile cellular subscriptions per 100 inhab.**		**International Internet bandwidth Bit/s per Internet user**		**Proportion of households with computer**		**Proportion of households with Internet**	
	2008	2010	2008	2010	2008	2010	2008	2010	2008	2010
Romania	22.0	20.9	113.3	114.7	29860	51408	37.8	47.9	30.4	42.2

Use indicators	**Percentage of individuals using the Internet**		**Fixed (wired)-broadband Internet subscriptions per 100 inhabitants**		**Active mobile broadband subscriptions per 100 inhabitants**	
	2008	2010	2008	2010	2008	2010
Romania	32.4	39.9	11.5	14.0	21.3	32.8

Skills indicators	**Gross enrolment ratio**				**Adult literacy rate**	
	Secondary		**Tertiary**			
	2008	2010	2008	2010	2008	2010
Romania	91.6	93.2	65.6	66.9	97.7	97.7

Source: ITU World Telecommunication/ICT Indicators database.

The Networked Readiness Index and the decile approach. In the Networked Readiness Index (NRI), the analysis of country performance is based on decile rankings. A decile ranking attributes ranks based on scores while taking into account the number of countries in the sample. For each edition of the NRI, the overall ranking was divided into 10 segments called deciles, each with an equal count of ranks. The 1st and 10th deciles comprise the economies that rank the highest and the lowest, respectively. Based on this approach, the 50th rank may correspond to various decile in two different moments of time.

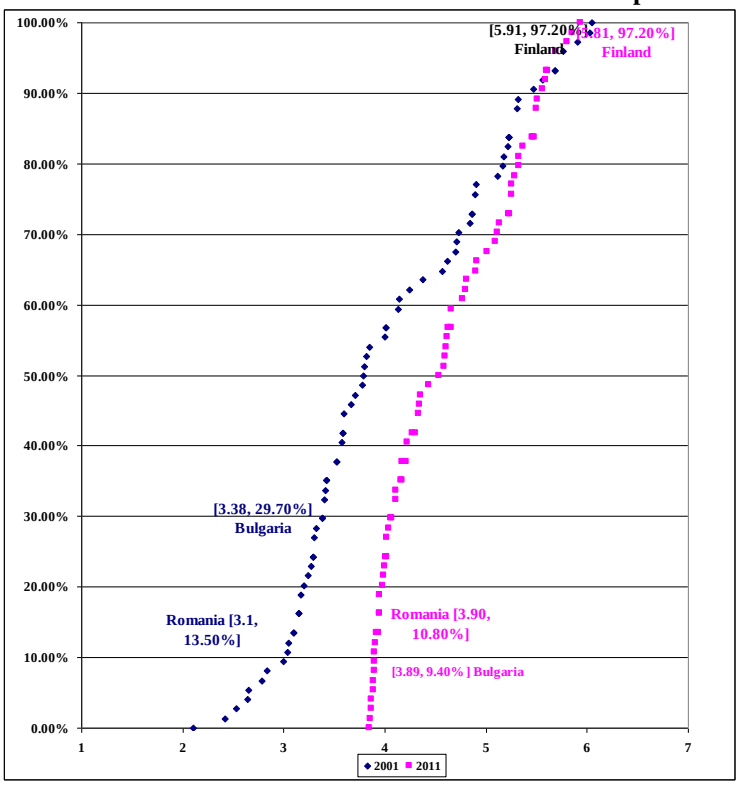
The 2001-2002 edition of the Networked Readiness Index (NRI) has been used to assess the comparative progress of eighty-two countries along different dimensions of progress in ICT; the last report compared 142 countries. This allows a comparison of countries' performances over time in the presence of varying sample sizes. In Europe, Romania was among the worst performers in 2001–2002 (ranked 65th out of 75 countries); in the 2011-2012 edition of the NRI, Romania belongs to the top half of the ranking reaching a significant gain of four decile ranks – table 3. It is followed by Bulgaria with a gain of two decile ranks, and, as it was expected the other two countries kept their superior performance – being in the first performance decile.

Table 3. Evolution in decile rankings since first inclusion

Country	Previous inclusion 2001-2002		Current edition 2011-2012		Difference gain (-) or loss (+)
	Rank/score	**Decile**	**Rank/score**	**Decile**	
Sweden	4th/5.76	1	1st/5.94	1	
Finland	3rd/5.91	1	3rd/5.81	1	
Romania	65th/3.10	9	67th/3.90	5	+4
Bulgaria	53rd/3.38	7	69th/3.89	5	+2

For further analysis, the cumulative distribution function of the indicator values was used to picture the distribution of the 2001 and 2011 distribution of performance in the corresponding samples of countries – figure 1. Also, we are interested in the difference between the two decile ranks, as an indication of the economy's dynamism – table 3. The estimate of the cumulative probability is intuitively the percentage of the data points to the left of the point of interest. In figure 1, Romania present only 13,50% on the performance profile in 2001, lowering to 10,80% in 2011 – each time compared to the rest of the countries included in the sample. Bulgaria reveals a similar evolution, and Finland and Sweden are in top of the ranking.

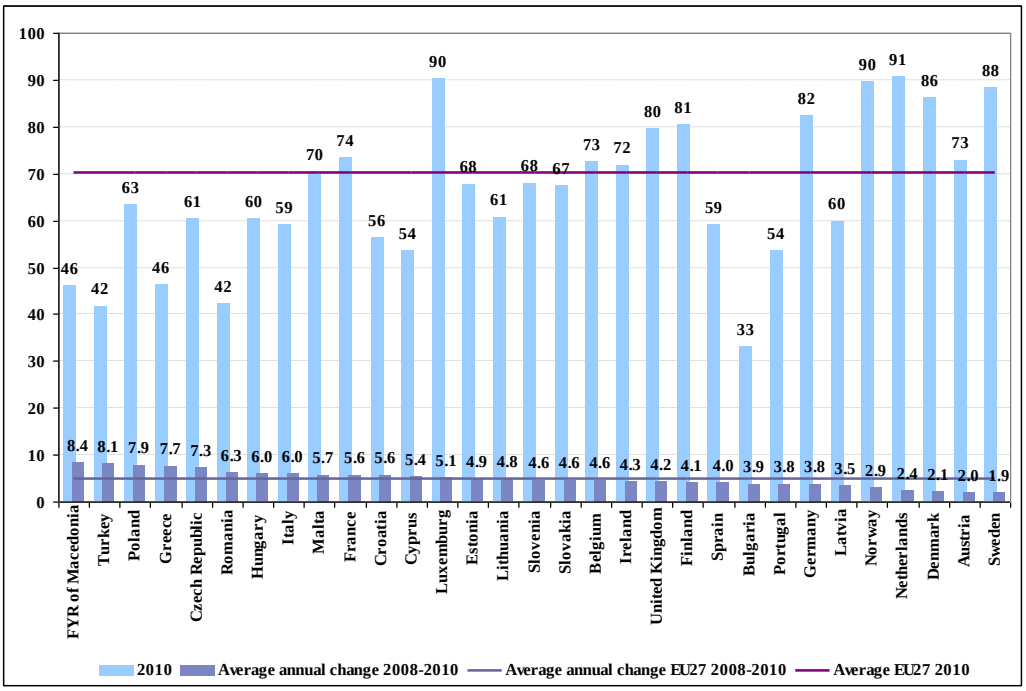
Figure 1. The cumulative distribution function for the NRI performance



4. The comparative analysis based on the EUROSTAT data

To complete the picture of the precarious position of Romania the EUROSTAT data on the information society from 2011.

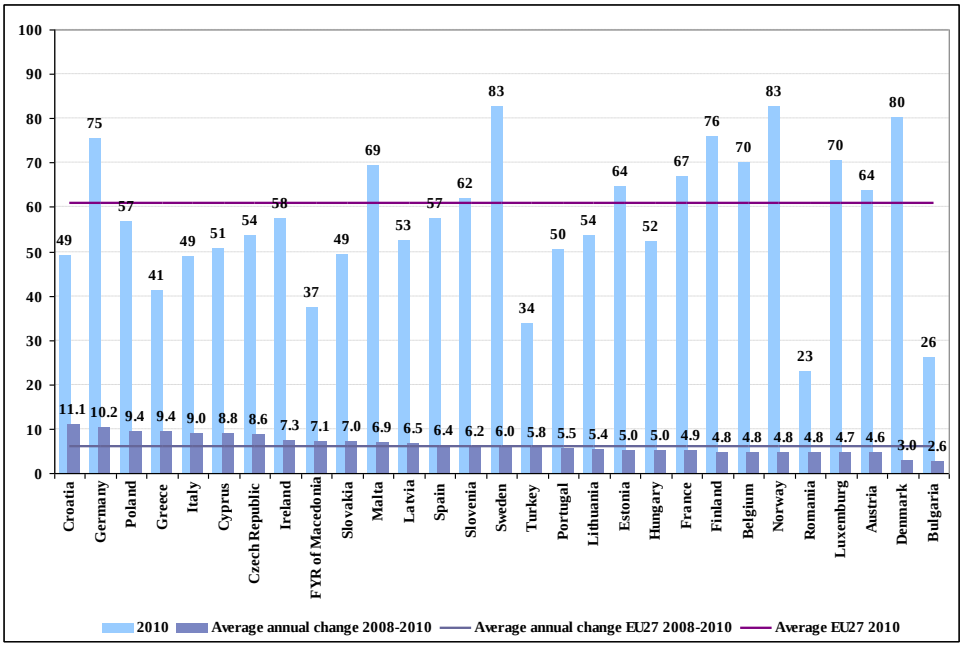
Figure 2. Internet access in households 2008 - 2010



Source: Eurostat (online data code: isoc_ci_in_h)

The Internet access is expressed in figure 3 as the share of households with internet access in 2010; the average annual change, in percentage points is represented, as well. Bulgaria reaches the 33%, Romania 43%, very low levels as compared to those of Finland of 81% and Sweden of 88%. Neither in terms of annual increase, the numbers are not very comforting – yet, Romania grew with 6,3% in 2008-2010, Bulgaria registered a growth of 3,9%, Finland of 4,1% and Sweden has reached the critical mass of internet users – it report only 1,89%.

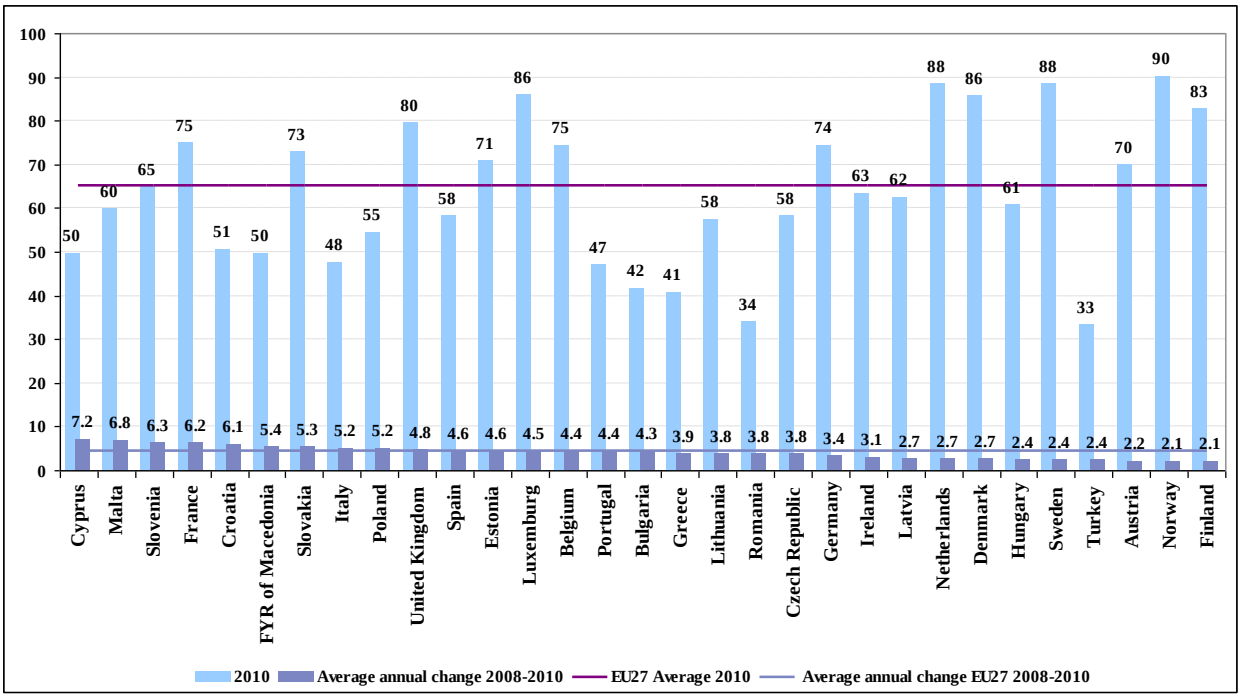
Figure 3. Broadband connections in households 2008 - 2010



Source: Eurostat (online data code: isoc_si_broad)

In figure 3, there is presented the distribution of broadband connections defined as the share of households with broadband connection in 2010 and average annual change, in percentage points. It may be seen that Romania is registering only 23 p.p., quite a low level as compared to the average level for EU27 states with a value of 57%. Even in terms of annual growth, the figures are also unfavourable Romania reporting an increase of only 4.8% lower that the average annual change in the EU27 of 4.5%.

Figure 4. Regular use of the Internet 2008 - 2010



Eurostat (online data code: isoc_ci_ifp_fu)

In terms of the share of persons who accessed the Internet in 2010, on average, at least once a week and average annual change, in percentage points), Romania is placed in the lower level (with 34% of the respondents stating about regular Internet use), Bulgaria reports 42%; meanwhile, Finland and Sweden reveal one of the highest value: 83% and, respectively, 88%.

5. The Eurostat indicators on skills issues

Since 2003, Eurostat has collected annual statistics for the EU and EFTA countries on access to, and use of, ICTs by individuals, households and enterprises. So far data collection to measure progress in the information society has been structured around the issues of take up of available connectivity opportunities, the use of ICT and possible obstacles, e-Inclusion and e-Skills, the diffusion of eCommerce and eServices. It covers households, businesses and the public sector. Current indicators will need to be adapted to reflect the emergence of new technologies and their usage. However, the issue of burden for respondents and National Statistical Institutes should be taken into consideration.

The Eurostat surveys on ICT use illustrate how rapidly Internet usage has increased throughout Europe. Yet, there is still a lack of assessment reports in dealing with digital skills. There are two sets of skills surveyed in the Eurostat surveys: computer related and internet related skills. Digital literacy is mainly measured through a list of tasks performed with a computer and when using the internet, from which the following indicators are derived: percentage of individuals with computer skills (none, low, medium, high) and % of individuals with internet skills (none, low, medium, high). The indicators are biannually updated and extended to cover new usages, mobile access to the internet, web 2.0 technologies etc. and should be linked to education.

The Eurostat registers data on "*E-skills of individuals and ICT competence in enterprises*" using two indicators "Individuals' level of *computer* skills" (in terms of percentage of the total number of individuals aged 16 to 74) and "Individuals' level of *Internet* skills" (also, as percentage of the total number of individuals aged 16 to 74). The first indicator "Individuals' level of computer skills" is defined in three versions having associated three levels:

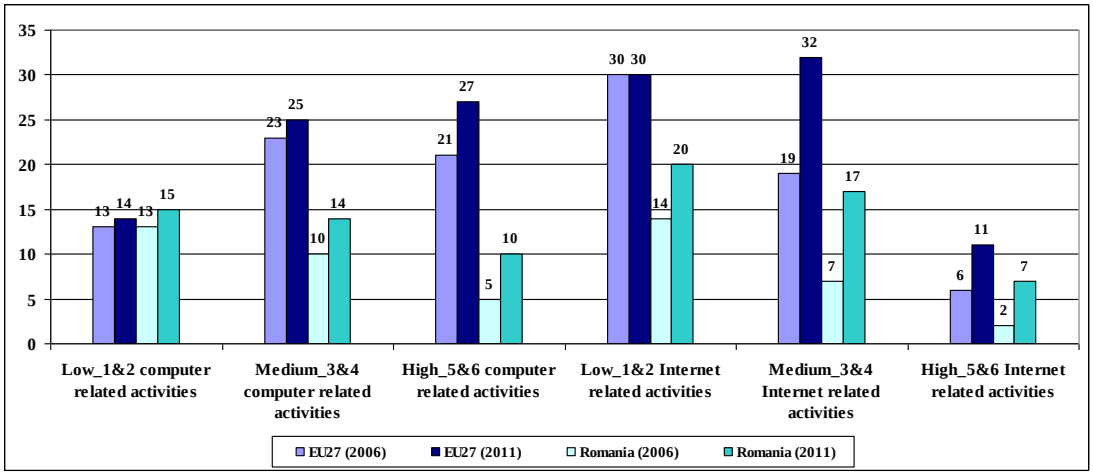
- low level of basic computer skills: corresponding to those individuals who have carried out 1 or 2 of the 6 computer-related items;
- medium level of basic computer sills: corresponding to those individuals performing 3 or 4 of the out 6 computer-related items
- high level of basic computer skills: counting the individuals who have carried out 5 or 6 of the 6 computer-related items.

Each level of basic computer skills is measured using a self-assessment approach, where one respondent is asked to indicate whether he/she has carried out specific tasks related to computer use (without these skills being assessed, tested or actually observed). Six computer-related items were used to group the respondents into levels of computer skills: copy or move a file or folder; use copy and paste tools to duplicate or move information within a document; use basis arithmetic formula (add, subtract, multiply, divide) in a spreadsheet; compress files; connect and install new devices, e.g. a printer or a modem; write a computer program using a specialized programming language. Instead of the item on having connected and installed new devices, the 2005 items included the use of a mouse to launch programs such as an Internet browser or word processor.

Almost the same kind of measurement applies for the other indicator that is split into: low level, medium level and high level of basic internet skills. Level of internet skills are measured using a self-assessment approach, where the respondent indicates whether he/she has carried out specific tasks related to internet use, without these skills being assessed, tested or actually observed. Six Internet-related items were used to group the respondents into levels of Internet skills: use a search engine to find information; send an e-mail with attached files; post messages to chat-rooms, newsgroups or any online discussion forum; use the Internet to make telephone calls; use peer-to-peer file sharing for exchanging movies, music etc.; create a web page.

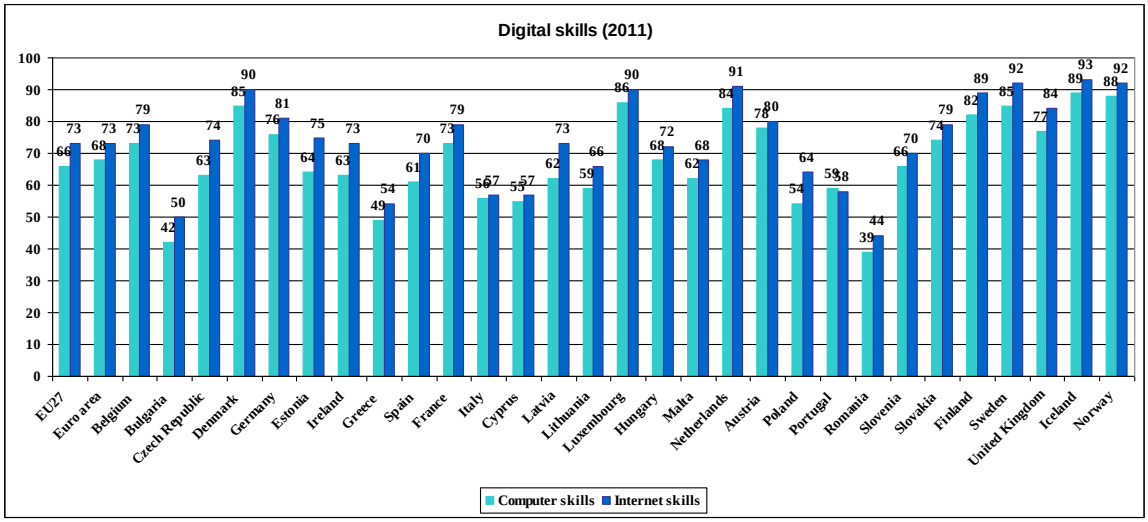
Regardless of how the skills divide is measured, starting from a commonly agreed form of definition, there is a significant disparity in population's digital skills among various European state members. This will be the point of interest in the following statistical analysis, attempting to reveal the size of the digital and skill divide for Romania, as compared to Bulgaria, quite similar in the geopolitical context and very dissimilar to Sweden and Finland – the leaders in the ICT – driven society. The relevance of this gap is considered as crucial important as recently there is a recognized that recent developments in digital skills acquisition have been very heterogeneously spread across Europe.

Figure 5. The distribution of the digital skills by levels



Source: e-Skills [isoc_bde15csk]

Figure 6. Percentage of individuals with computer and internet skills (2011) – all levels cumulated



Source: e-Skills [isoc_bde15csk]

In line with the rise in internet use in Europe, more and more Europeans are acquiring digital skills. In 2011, 66 % of Europeans had at least some level of computer skills and 73% of Internet skills. However, this overall picture masks a very diverse landscape of digital skills distribution in Europe. In Sweden, Netherlands, Denmark, Iceland and Norway, over 90 % of people possess digital (both computer and Internet) skills. At the same time, approximately, 39% of Romanians and 49% of Bulgarians possess computer skills; similarly, 44% of Romanians and 50% of Bulgarians possess Internet skills, making them, along with Greeks citizens almost unable to fully and successfully participate in the digital age.

6. The skills acquiring and widening the gap for vulnerable groups

There are still significant disparities in digital skills not only among European countries, but also among different socio-economic groups. However, the Riga Indicator for digital skills rose by 0.02 in 2009 compared to 2007. The most digitally skilled groups among disadvantaged citizens include women (0.95), the unemployed (0.92) and people living in lagging (0.80) or rural (0.86) areas. Significantly lagging behind are senior citizens aged 65 to 74 (0.38) and economically inactive people (0.52) (Europe's Digital Competitiveness Report, 2010).

When considering the possible disadvantaged population groups individually, it is expected that the population groups more distanced (i.e. with the larger disparities) from the EU regular digital skills' level are the aged over 55 years old, the population living in sparse density areas, and placed in terms of household's incomes in the first two quartiles. For example, underperformance in terms of digital skills among disadvantaged groups, in particular older people, those with a low education or those on low incomes, puts them at a disadvantage in the emerging digital society. In particular, these divides are significantly deepening with the increasing sophistication of digital skills.

As the digital literacy remains a subject of challenge for the policy administrators for both supply and demand sides, more efforts need to be dedicated to supporting potentially deprived or vulnerable groups. From the literature and even empirical observations, it can be noticed that in terms of the skills' possession rates of the specific disadvantaged groups, most countries have the largest disparities in Internet use with the low educated, the economically inactive, and the aged 65-74.

Policy concentrates on at-risk groups to prevent that disadvantaged people and disadvantaged groups are left behind in the information society. A key factor is to enable all citizens to keep up with technological developments that affect their daily life and their employment prospects. Disadvantages can be related to differentials in access, bandwidth, skills and digital media literacy, service usage and quality of usage of new services and information and communication technologies.

At-risk groups and geographical breakdowns include the following: older people, women, people with low education levels, unemployed, people with disabilities, marginalised young people, income groups, immigrants, ethnic groups, people living in less developed regions, people living in rural settlement types, period of use.

This paper starts from the Internet disparity index (introduced in 2006) and uses recent Eurostat data (the majority from 2011 and, in the case of segmentation for computer skills from 2009, as exception). This index was originally been produced to measure the disparities between the disadvantaged groups and the European average as collected by Eurostat. The overall index used by authors to assess the current state of computer and Internet skills is based on penetration rate ratios of various types of digital skills (in terms of percentage of all individuals), which has been shown to be the most appropriate way to track and analyze disparities over time the digital skill gap.

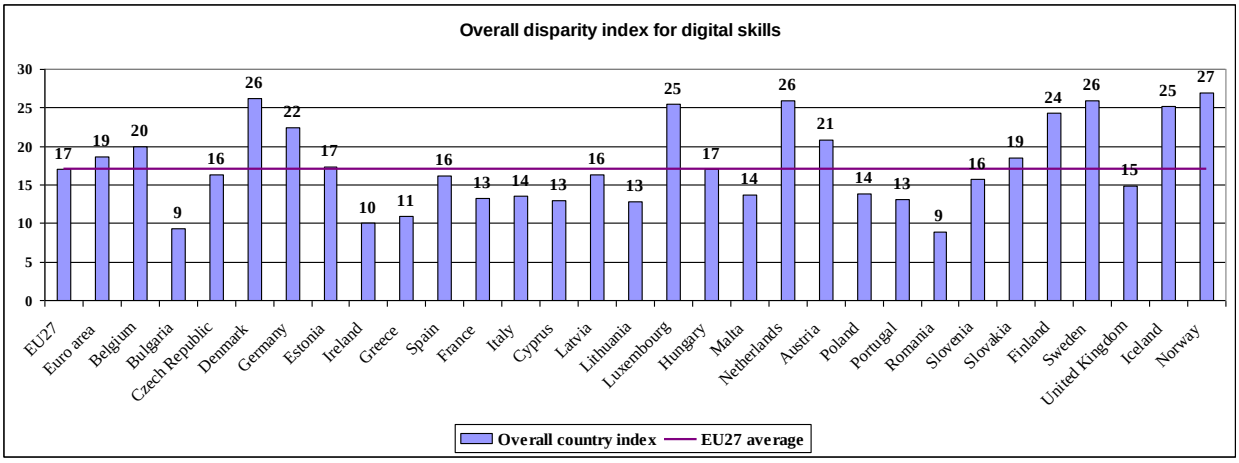
Table 4. Index of digital skills (regardless the level) in at risk groups by country in 2009-2011

	Computer related activities			**Internet related activities**		
	Age segmentation (2009)	**Population density segmentation (2011)**	**Income quartile segmentation (2011)**	**Age segmentation (2010)**	**Population density segmentation (2011)**	**Income quartile segmentation (2011)**
EU27	12	19	17	14	19	23
Euro area	12	21	18	14	24	24
Belgium	11	23	18	16	26	26
Bulgaria	3	9	8	4	16	15
Czech Rep.	8	19	12	12	25	22
Denmark	20	28	27	23	30	30
Germany	18	24	22	18	27	26
Estonia	6	20	18	11	25	25
Ireland	8	18	0	10	24	0
Greece	3	12	11	4	19	18
Spain	7	18	16	8	23	24
France	15	22	0	18	25	0
Italy	7	17	13	8	19	18
Cyprus	5	15	14	5	19	20
Latvia	5	20	16	9	25	23
Lithuania	5	17	12	8	15	21
Luxembourg	22	29	21	24	30	27
Hungary	10	20	16	11	24	22
Malta	7	8	16	8	20	23
Netherlands	20	27	25	23	30	30
Austria	12	25	22	14	26	26
Poland	5	16	13	7	22	20
Portugal	6	15	12	7	19	19
Romania	3	9	7	4	16	13
Slovenia	7	19	13	9	24	21
Slovakia	9	24	21	13	18	27
Finland	15	27	25	20	30	29
Sweden	18	27	25	24	31	30
United Kingd.	15	27	0	19	28	0
Iceland	18	29	29	23	21	32
Norway	22	30	26	23	30	31

Source: Based on Eurostat 2011 ICT Community Survey of Household and individuals (includes all individuals)

A number of countries have an at-risks groups index value much lower than the EU27 17% value. In particular, in the EU countries, Bulgaria and Romania (with 9%), Ireland (with 10%) and Greece (with 11%) has lower values below 17% (meaning that the at-risk groups in these countries posses digital skills in percentages smaller than quarter of the total population in each country). The EU country with the highest value (i.e. the lowest disparity) is Sweden at 26% and from the EFTA states is Norway with 27%, meaning that in Sweden and in Norway at-risk groups posses digital skills at approximately 26-27% of the whole population.

Figure 7. The regional distribution of the digital skills vulnerability index (2011)



The z-score approach. The methodology is based on the computations of the Z-scores which are used to compare performance on diverse dimensions on a relative basis. A Z-score measures the number of SD units a value deviates from the mean (the number of SD units a given score X deviates above or below the mean score), it used to convert scores to percentile rankings and to determine the area or percentage of scores under the normal curve that correspond to a given range of Z-scores: $Z = \frac{X - \overline{X}}{SD}$ (1).

The figure 8 reports the distances from the EU average (measured in standard deviations); data are presented in such a way that data bars pointing to the right (left) always indicate performance which is better (weaker) than the EU average. It is most indicative in signalling the disparities of Romania regarding the computer skills (-0.84) and Internet skills (-0.80); the supremacy of Finland and Sweden is obvious.

Figure 8. The digital skills and various category of Internet usage

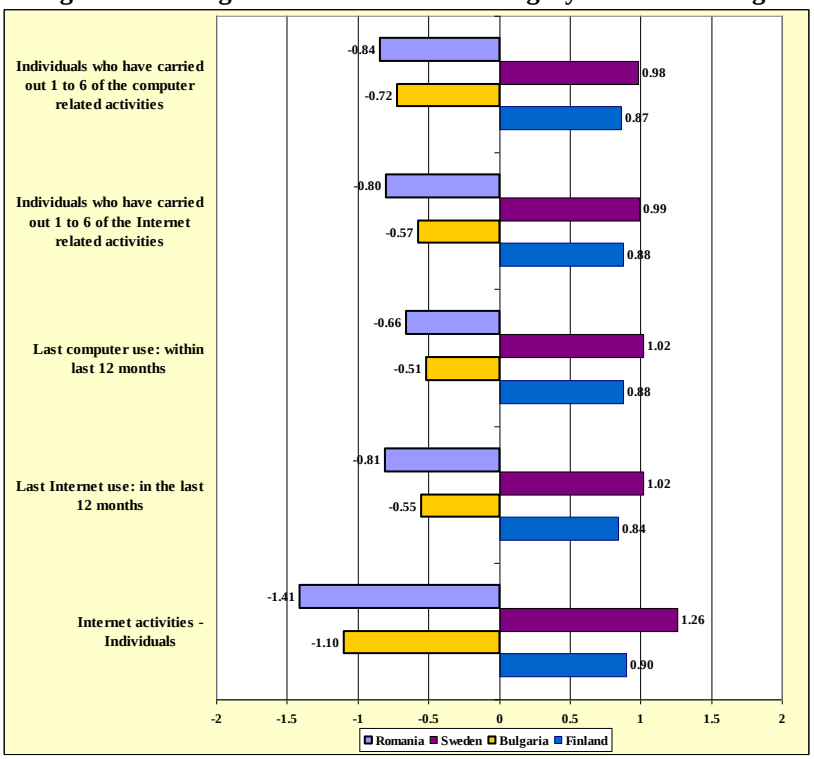


Table 5. The digital skills and various category of Internet usage

	Individuals with 1 to 6 of the computer related activities	Individuals with 1 to 6 of the Internet related activities	Last computer use: within last 12 months	Last Internet use: in the last 12 months	Internet activities - Individuals
EU27	66	73	75	73	63
Bulgaria	42	50	52	51	39
Romania	39	44	48	44	33
Finland	82	89	90	89	78
Sweden	85	92	94	94	85
Mean indicator	60.17	65.42	67.22	66.02	60.42
Standard Deviation	25.21	26.88	27.50	27.34	19.48

7. Conclusions

Digital Literacy is a crucial and challenging area for policy makers to achieve the targeted performance offered to all citizens by the information and knowledge society. All countries have put in place initiatives addressing the digital literacy of the general public, and most countries have also introduced schemes targeting specific disadvantaged groups. Despite its inclusive effects, the information society has also created new divides and one that inspired this paper deals with monitoring the digital skills. Indexes are often used to monitor policy measures.

Nowadays, the e-skills' possession allows a more effective participation in the global information economy and society, spreading the opportunities to conduct successful business and entrepreneurial endeavour or more simply facilitates active civil life, accessing public e-services.

Acknowledgment: This work was supported by the project "Post-Doctoral Studies in Economics: training program for elite researchers - SPODE" co-funded from the European Social Fund through the Development of Human Resources Operational Programme 2007-2013, contract no. POSDRU/89/1.5/S/61755.)".

8. References

- Van Deursen, A.J.A.M. & Van Dijk, J.A.G.M. (2009). Using the Internet: Skill Related Problems in Users' Online Behavior. *Interacting with Computers, 21*, 393-402
- Hîncu D., Frăţilă L., Tanţău A. Gap Indicator For Measuring Digital Divide, Management Research And Practice, Vol. 3 Issue 2 (2011) pp: 74-88
- Hüsing, T., Hannes S., (2003): A Digital Divide Index – an Approach for the measurement and benchmarking of social inequality in the diffusion of ICT in IT Diffusion, IT & Society, Issue 7, 21-38
- Kirkman, G.S., Osorio, C.A., Sachs J.D., (2002) The Networked Readiness Index: Measuring the Preparedness of Nations for the Networked World, chapter 2, http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan008655.pdf
- Mason, S.M., K.L. Hacker (2003): Applying Communication Theory to Digital Divide Research, IT & Society, Issue 5, 40-55
- Measuring the Information Society 2011, International Telecommunication Union© 2011 ITU
- Benchmarking Digital Europe 2011-2015 - A conceptual framework, Issue No: 27, October, 2009
- Digital Literacy Report: a review for the i2010 e-Inclusion initiative, Digital Literacy European Commission Working Paper and Recommendations from Digital Literacy High-Level Expert Group
- Computer skills in the EU27 in figures, EUROSTAT news release 47/2012 - 26 March 2012

- Digital Agenda Scoreboard 2011, Directorate General Information Society of the European Commission: http://ec.europa.eu/information_society/digital-agenda/scoreboard/index_en.htm; the pillar "Digital Competence in the Digital Agenda"
- Emerging Skills and Competences- A transatlantic study EU-US Study for the European Commission, October 2011 http://ec.europa.eu/education/more-information/doc/2011/skills_en.pdf
- Eurostat (2012), Structural indicators, http://epp.eurostat.ec.europa.eu/portal/page/portal/structural_indicators/indicators
- Europe's Digital Competitiveness Report, COMMISSION STAFF WORKING DOCUMENT, Vol. I, Brussels, 17.5.2010 SEC(2010) 627
- Europe's Digital Competitiveness Report, ICT Country Profiles, Vol. II, COMMISSION STAFF WORKING DOCUMENT Brussels, 17.5.2010, SEC(2010) 627
- Measuring progress in e-Inclusion, Riga Dashboard, 2007, http://ec.europa.eu/information_society/activities/einclusion/docs/i2010_initiative/rigadashboard.doc
- Romania profile; http://ec.europa.eu/information_society/digital-agenda/documents/countryprofiles.pdf

A RESEARCH ON RETRIEVING AND PARSING OF MULTIPLE WEB PAGES FOR STORING THEM IN LARGE DATABASES

BUCUR Cristian
Teaching assistant, Faculty of Economical Sciences, Department of Modeling, Economic Analysis and Statistics, Petroleum-Gas University of Ploiesti, Ploiesti, Romania, Ph.D. student ISUD ASE Bucuresti, Romania, cr.bucur@gmail.com

TUDORICA Bogdan George
Teaching assistant, Faculty of Economical Sciences, Department of Modeling, Economic Analysis and Statistics, Petroleum-Gas University of Ploiesti, Ploiesti, Romania, Ph.D. student ISUD ASE Bucuresti, Romania, tudorica_bogdan@yahoo.com

Abstract: *This paper intends to present one of the studies we jointly done during the research for our Ph.D. theses. Cristian Bucur's thesis aim is to study how the knowledge stored in web pages from various sources can be retrieved and classified. Bogdan Tudorica's thesis aim is to study the ways to manage large quantities of data for various purposes (especially through use of new technologies, such as NoSQL databases. As such, the application we are describing in this paper is a mixed one, containing both web page crawling and parsing and data storage in a commonly used NoSQL database.*

Key words: web page crawling, web page retrieval, web page parsing, NoSQL database

JEL classification: C83, C88, C89

1. Introduction

As we mentioned in the abstract, the main target of this paper is to describe an application we produced as a component for our PhD thesis. The application in itself, at this stage, is more of a technological demonstration than a production application, lacking both the performance and the capabilities required for a complete effective product. As such we purposely used in our application technologies and components which will not ensure optimal performance but instead will be easy to use and especially will create a clear picture of what is needed for consistently finding and extracting specific data from multiple web sources followed by storing that data in a container system able to handle large quantities of data with the structure which is specific for web pages content. To be more precise, we specifically targeted web sites which are hosting online publications, to give for a certain type of content, structured in articles.

2. Technical details of the implementation

The application in itself is written in the PHP language and consists of 4 files called test_mongo.php (which is starting the application and providing user interaction), scraper_mongo.php (which is "scraping" the initial and the subsequent web pages, collecting the URL addresses and storing them in the database; scraper_mongo.php was initially proposed by Aziz S. Hussain under a GPL license and heavily modified by us for our application), parse.php (which is parsing the content obtained from the web pages, eliminating the HTML tags and replacing the codes used in HTML for special characters with their equivalent chars) and mongo_admin.php (which is a simple MongoDB administration interface proposed by Jonathan H. Wage; this interface was designed to be called from an external application, such as ours; we used this interface as it was, being perfectly suited to our purpose). For our implementation we used the 5.3 version of the PHP environment.

The choice of using PHP was a weighed one, having both advantages and disadvantages. The advantages are related to portability (our application can run on any of the many PHP enabled web servers available at this moment) and simplicity. The disadvantages are mainly given by the reduced performance which can be obtained by running an interpreted PHP script compared with the performance which can be virtually obtained by using a local compiled application which is also specifically optimized for the specifics of the hardware and operating system used. The fact that PHP is not capable in itself to provide multi-threading capabilities (although there are available some multi-threading implementations

of PHP which are based on the CURL extension or some other means – but all of them are making the implementation more complicated) is another performance reducing detail.

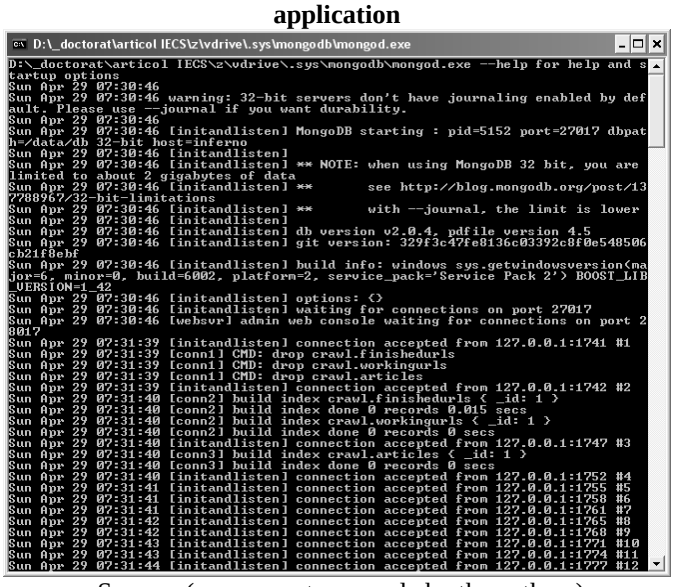
The database system used for our application is MongoDB which, unlike PHP, is a perfect choice even for a better performing application. MongoDB is an open source, document-oriented database designed to store binary JSON documents with dynamic schemas. This new data model simplifies coding significantly, and also improves performance by grouping relevant data together internally.

The advantages given by MongoDB which are making it such a good choice are given by:

- Its high portability (it's open source and it has pre-made builds for OS X 32-bit, OS X 64-bit, various distributions of Linux 32-bit and Linux 64-bit, Windows 32-bit, Windows 64-bit, Solaris i86pc, Solaris 64; the fact that it's also available as source is making it easily available for other operating systems such as BSD, BeOS and various flavors of UNIX).
- Its ability to store large amounts of data (although its 32-bit versions are limited to 2GB storage space).
- Its high performance (the reason for this is that MongoDB uses memory-mapped files for performance – as such, the speed limits of the storage devices are not influencing the performance of the database system).
- The large number of drivers available for accessing a MongoDB database (C, C#, C++, Erlang, Haskell, Javascript, Java, Perl, PHP, Python, Ruby, Scala).
- MongoDB is neither a relational database (which have rich functionality) nor a key-value stores type database (which are fast and scalable) but something in between. As such MongoDB maintains many of the great features of a relational database, like indexes and dynamic queries but also has all the advantages of the key-value stores (speed, ease of operation and of use, high horizontal scalability).

For our application we used the 32-bit version of MongoDB (the amount of data we intended to extract was well within the 2GB limit we mentioned before and we were forced to do so by the fact the three machines we used to program and test the application had 32-bit operating system installed on them). Figure 1 shows the start-up sequence of the MongoDB server used for our implementation, followed by the journal entries of several connections made by our application to the database server.

Figure 1: The start-up sequence of the MongoDB 32-bit server, followed by the first connections made by the application



Source: (screen capture, made by the authors)

As for the working environment of the application, the operating system was Windows (Windows XP Professional 32-bit edition on two the machines we used for programming and testing and Windows 7 32-bit edition on a third one). The server package we used for hosting the application was ZWAMP, a lightweight zero-install Web server package that runs on Windows.

We used ZWAMP version 2.1.1 which contains Apache 2.4.2, MySQL 5.5.22, PHP 5.3.10, MiniPerl 5.14.2, MongoDB 2.0.4 (PHP 1.2.10 driver), APC 3.1.8dev, eAccelerator 1.0svn427, MemCached 1.4.5, XCache 1.3.2, XDebug 2.1.3, Adminer 3.3.4 and PHP MongoDB Admin. The choice of using ZWAMP was a subjective one, based on our previous experiments with the various server

packages available (WAMP, XAMPP, EasyPHP, Uniserver, etc.) and it aimed for a higher stability and simplicity of the implementation. As such, our application can presumably run on any other of the previously mentioned server packages, provided that the package has the same PHP version (for avoiding inter-version incompatibility) and has the needed MongoDB driver.

We have to mention that we made multiple modifications to the php.ini configuration file for making the server package usable for our implementation:

- Added character set setting to provide for regional variations: default_charset="UTF-8"
- Changed memory limit to accommodate for larger data quantities: memory_limit=100M
- Changed error reporting for debugging: error_reporting=E_ALL | E_STRICT
- Changed error displaying for debugging: display_errors=On
- Changed variable order of registering: variables_order="GPCS"
- Added variable order of requesting: request_order="GP"
- Added parameter for blocking the usage of long arrays (for increased security of the application): register_long_arrays=Off
- Blocked the usage of argv and argc variables (for increased performance): register_argc_argv=Off
- Limited the PATH variable to the current path only: include_path="."
- Limited the number of simultaneous uploaded files: max_file_uploads=20
- Blocked the inclusion of URL addresses: allow_url_include=Off
- Added the use of several drivers and libraries: extension=php_igbinary.dll; extension=php_xcache.dll; ;extension=php_xdebug.dll
- Accelerated garbage collection: session.gc_divisor=1000
- Limited the compatibility with the previous PHP version: session.bug_compat_42=Off
- Various other modifications and added parameters: apc.shm_size=16M; apc.max_file_size=4M; xcache.var_size=8M

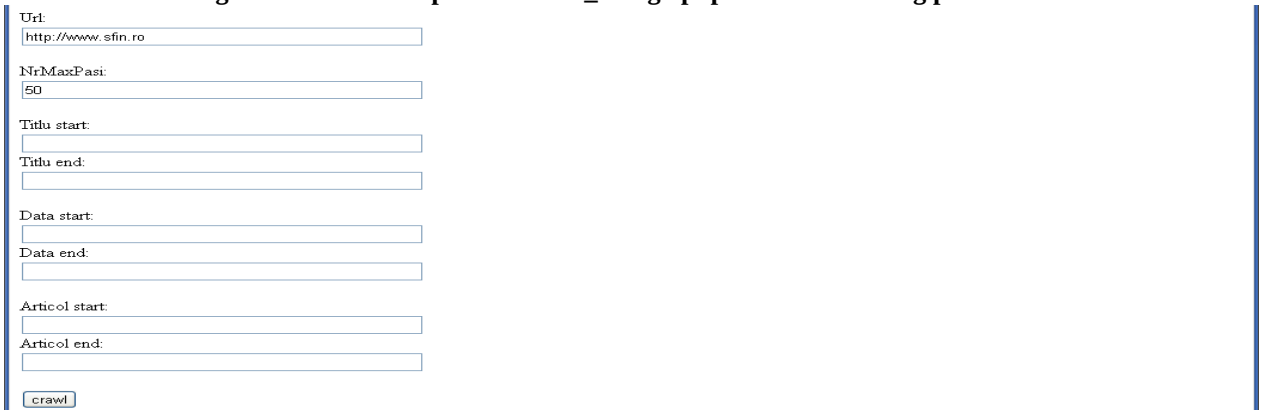
The user interface of the application consists of two components. The first one is a simple form with several text fields in which the user can type the running parameters (the URL of the main page which will be "scrapped", the maximum number of parsed pages and several selection criteria: by title, by date and by article, all of them given as intervals – starting and ending value).

All the text fields have visible or invisible default values (http://www.sfin.ro for the default starting URL – this one is a Romanian online economical publication with a large number of articles called "Saptamana Financiara" – The Financial Week, and as such it's a pretty good candidate for an application run; 4 for the default number of analyzed pages – to make for a fast demo run; first possible and last possible value for the default filtering details).

The code used for applying the parameters given by the user is the following: function setDelimiters($maxsteps=3,$ts,$te,$ds,$de,$as,$ae) {// Set start URL; $this->maxSteps = $maxsteps; $this->timeS = $ts; $this->timeE = $te; $this->dateS = $ds; $this->dateE = $de; $this->artS = $as; $this->artE = $ae; }. This code and the rest of the code which ensures user interaction and the calling subsequent actions is contained in the test_mongo.php file mentioned at the beginning of the part 2 of this paper.

This user interaction form can be seen in the Figure 2.

Figure 2: The start-up form of test_mongo.php with the working parameters

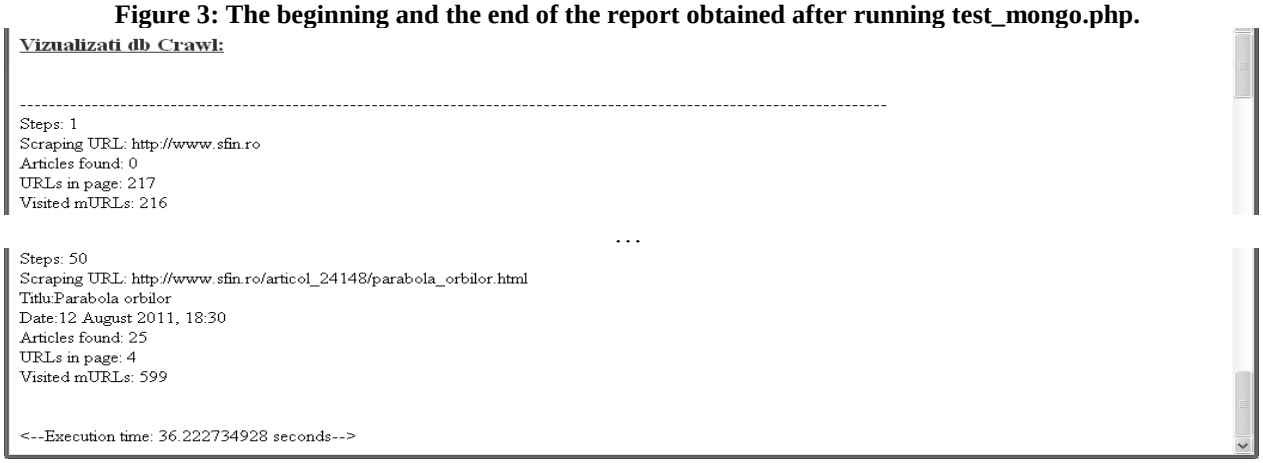


Source: (screen capture, made by the authors)

The same window used for the user interface is reused after running for displaying a log of the execution. Each address which was found and followed is listed as URL address, together with the number of articles found to that point, the number of subsequent URL addresses to that page and the number of visited URL addresses from the total. Again, for the sake of simplicity and performance, we limited the depth to which the application is going in the URL "tree" to one level only (which is good enough, taking into account the fact that on most online publications, all the articles are at the same level).

At the end of the journal, the running time for the application is given, measured in seconds and fractions of a second. Figure 3 shows the beginning and the end of a typical execution journal, for a limit of 50 analyzed pages, without any explicit content filter (the default values are used). As the second half of the Figure 3 shows, the execution time for these parameters is around 35-36 seconds. For 125 analyzed pages, the execution time is around 79-80 seconds.

Figure 3: The beginning and the end of the report obtained after running test_mongo.php.



Source: (screen capture, made by the authors)

A more extensive analyze on the factors which are influencing the performance of such an application is the subject of a work in progress made by us for our doctoral researches. We can tell for the moment that a lot of factors are influencing the speed of web page retrieval and parsing but the most influent of these factors are the speed of the connection to the internet and the capability to run the application as a multi-threading task.

We must note that at the beginning of each application run the database is emptied, as such as to obtain clear demonstrative results for the execution (the code used for that is: $m = new Mongo(); $db = $m->crawl; $c = $db->finishedurls->drop(); $c = $db->workingurls->drop(); $c = $db->articles->drop();).

The crawling in itself is done using the following code:

```
function getContents($url)
{ $ch = curl_init(); // initialize curl handle
curl_setopt($ch, CURLOPT_HEADER, 0);
curl_setopt($ch, CURLOPT_VERBOSE, 0);
curl_setopt($ch, CURLOPT_USERAGENT, "Mozilla/4.0 (compatible;)");
curl_setopt($ch, CURLOPT_AUTOREFERER, false);
curl_setopt($ch, CURLOPT_CONNECTTIMEOUT,7);
curl_setopt($ch, CURLOPT_REFERER, 'http://'.$this->useURL);
curl_setopt($ch, CURLOPT_URL,$url); // set url to post to
curl_setopt($ch, CURLOPT_FAILONERROR, 1);
curl_setopt($ch, CURLOPT_FOLLOWLOCATION, 1);// allow redirects
curl_setopt($ch, CURLOPT_RETURNTRANSFER,1); // return into a variable
curl_setopt($ch, CURLOPT_TIMEOUT, 50); // times out after 50s
curl_setopt($ch, CURLOPT_POST, 0); // set POST method
$buffer = curl_exec($ch); // run the whole process
curl_close($ch);
```

return $buffer;}

The web page content parsing is ensured by the following code:

```
$title=parse($pageContent,$this->timeS,$this->timeE);
$date=parse($pageContent,$this->dateS,$this->dateE);
$article=parse($pageContent,$this->artS,$this->artE);
```

The data output to the database is done using the following code:

```
$table=$db->articles;
$article=iconv("UTF-8","UTF-8",$article);
$db->articles->insert(array("titlu" => $title, "data" => $date, "articol" => $article));
```

These code samples and the rest of the scraping and crawling process are contained in the scraper_mongo.php file mentioned at the beginning of the part 2 of this paper.

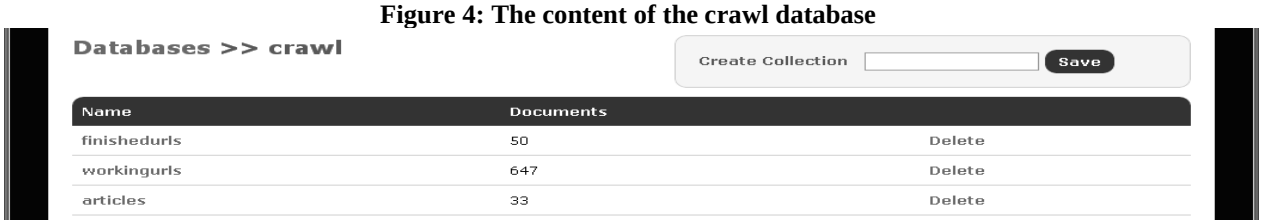
On the page used to journalize the execution, at the beginning, we placed a link to the MongoDB administration interface, targeted to the database used for our application, a database which is called "crawl".

As one can see in the Figure 4, the administration interface contains, for this database, the names of the collections contained in it (a collection is a MongoDB structure somehow similar to a table from a classical relation database; the differences as given by the fact that the content of a collection is not a set of instances of a relation but a set of key-value items tuples which are called documents).

We used a structure with three collections which will be further described. The Figure 4 contains the result of the database after a test run for a limit of 50 analyzed pages, without any explicit content filter (the default values are used). On the web site which was analyzed, only at the first level of links (see the earlier comments), 647 URL addresses were available, 50 of these were taken into consideration (this was the parameter we gave at run time), and from these 50, 33 linked to pages which contained articles, pages the application could parse.

We should note though, that taken into account the run times we obtained for values ranging from 50 to 125 pages and the fact that the default limit runtime for a php script is 90 seconds, we can expect that the application would fail for a number of 140-150 pages to be analyzed. This is not exactly a problem, taken into account the fact that we can easily override the limit from the php.ini file (we simply modify the parameter max_execution_time=90 to some bigger value). Also we already noted that this version of the application is only a technological demonstration, not a production application and such a quantity limitation is not really an issue.

Figure 4: The content of the crawl database

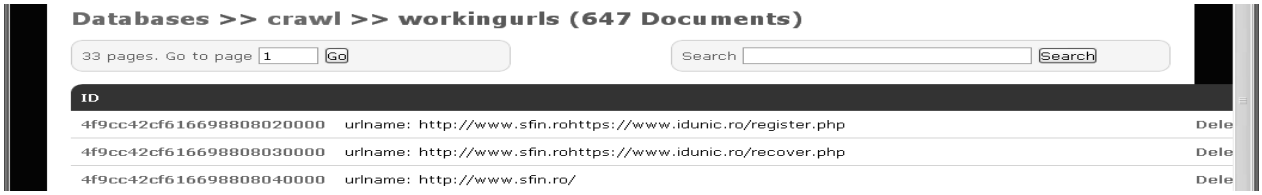


Source: (screen capture, made by the authors)

By following the three collection links available for the "crawl" database, one can browse through the URL addresses the application "scraped" on (the collection called "workingurls"), the URL addresses which were good candidates for parsing, and on which the application "crawled" (the collection called "finishedurls"), and finally the content and the metadata of the articles which our application extracted from these pages (the collection called "articles").

A sample of the content of the "workingurls" collection can be seen in the Figure 5. This collection has a simple structure with two keys, one ("_id") which is an auto-serialization key and one ("urlname") which contains the URL addresses which our application "scraped" from the starting page the user gave at the start time.

Figure 5: A part of the content of the "workingurls" collection



Source: (screen capture, made by the authors)

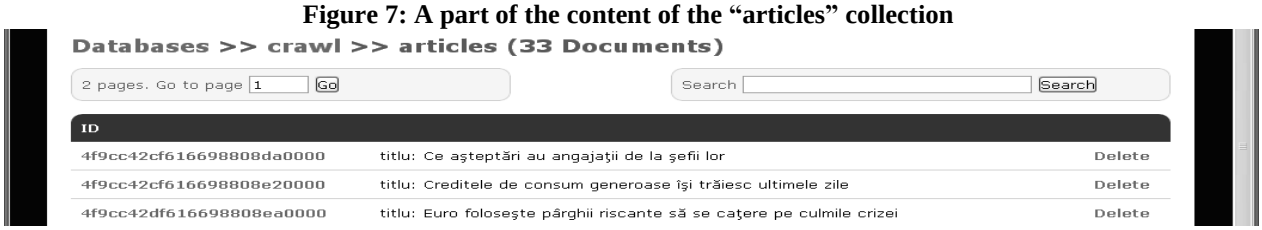
A sample of the content of the "finishedurls" collection can be seen in the Figure 6. This collection has a simple structure with two keys, one ("_id") which is an auto-serialization key and one ("urlname") which contains the URL addresses which our application "crawled" on, selected from the scraped pages collection. It's visible that the number of use pages from the site is way lower than the total number of pages from the site. We should expect as much, taking into account that the analyzed web site is not just an article store but contains many other types of information.

Figure 6: A part of the content of the "finishedurls" collection

Source: (screen capture, made by the authors)

A sample of the listing of the content of the "articles" collection can be seen in the Figure 7. This collection has a more complex structure, with four keys, first two of these being shown in the listing. First one ("_id") is an auto-serialization key and the second one ("titlu") contains the title of the article which was parsed from the article's web page.

Figure 7: A part of the content of the "articles" collection



Source: (screen capture, made by the authors)

The complete set of keys consists of four keys ("_id", "titlu", "data" and "articol") and a part of the content of such a key-value items tuple is shown in Figure 8. Again the "_id" key contains auto-serialization data.

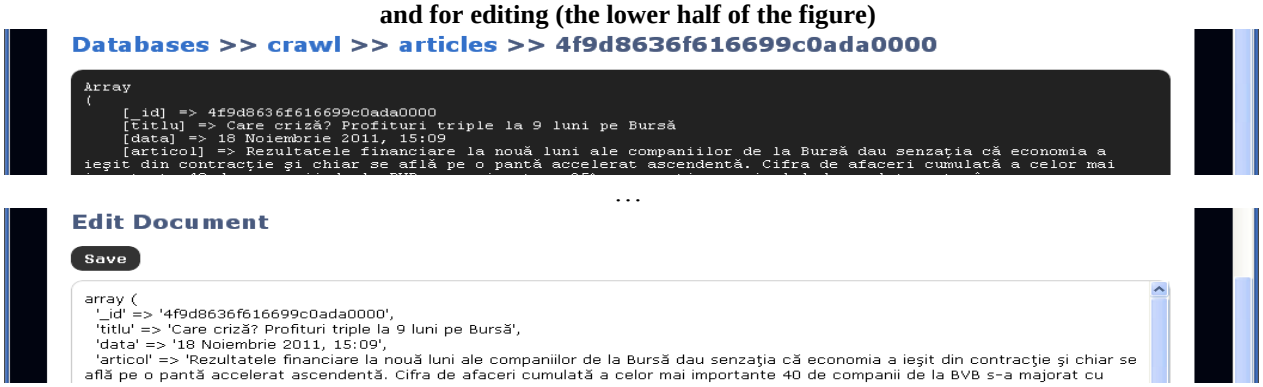
The "titlu" key contains the title for the article as it was parsed from the article's web page.

The "data" key contains the date and time at which the article was published, as it was parsed. It should be noted that data contained in this key (as any data contained in any key in a MongoDB database) is not really a date/time structure of any kind but a string (encoded UTF-8) which contains the date and time, exactly as it is written on the article's web page.

The "articol" key contains the text of the article as it was parsed from the article's web page. All the formatting information is removed, the only thing remaining being the text in itself. The code which cleans the formatting information is included in the parse.php file mentioned at the beginning of the part 2 of this paper.

As for the all other collections, the information stored in the "articles" collection can be directly edited from the MongoDB Administration interface (as it can be seen in the second half of the Figure 8) but this fact has little importance for our application (except for maybe correcting some data, before using it in a following application).

Figure 8: A sample article opened from the database for viewing (the upper half of the figure) and for editing (the lower half of the figure)



Source: (screen capture, made by the authors)

3. Future work and conclusions

The aim of this paper was to present one of the applications we produced during the research for our PhD thesis. As it was shown in the previous part of the paper, this application is capable of taking a given URL address which is supposed to be the address of an online publication, to crawl through the pages subsequent to the main page, to identify only those pages which are containing articles and to extract from these pages their content, in a certain fashion. The extracted data is then stored in a NoSQL MongoDB database.

As the application, although fully functional, is certainly lacking performance and is providing only a reduced set of features, is well under our attention to redesign the application for covering these topics. Possible improvements are:

- To change PHP with a more classical RAD environment which to natively offer multi-threading capabilities, code compilation and some optimization features.
- To add some flexibility regarding the web page regional code setting and the level of depth in the URL "tree" the application can reach.
- To parse additional metadata from the visited web pages.

Another line of work which we intend to follow in the near future is to design an environment and a framework for consistent performance measuring for our application and for similar ones.

Finally it should be noted that the byproduct of the application – the database containing the texts of the articles – was designed as such as it is ready to be used by another line of products from our research which is aiming at data analysis and classification (the "articol" key from the "articles" collection contains only clear text, without any formatting information; more than that, the text is automatically tagged with the data contained by the other keys).

In this paper we provided as much technical information about the application as possible without exceeding the limits given by the publication policy, including short excerpts from the code used to build the application. For more information, including the entire code we used, you can contact us at any time.

4. Acknowledgements

This article is a result of the project POSDRU/88/1.5./S/55287 „Doctoral Program in Economics at European Knowledge Standards (DOESEC)". This project is co-funded by the European Social Fund through The Sectorial Operational Program for Human Resources Development 2007-2013, coordinated by The Bucharest Academy of Economic Studies in partnership with West University of Timisoara.

5. References

- Aziz Saleh Hussain (2012), *A web developer blog*, available at http://www.azizsaleh.com/
- Bogdan George Tudorica, Cristian Bucur (2011), *A comparison between several NoSQL databases with comments and notes*, RoEduNet International Conference - Networking in Education and Research, 23-25 Iunie 2011, ISSN: 2068-1038

- Cesar D. Rodas (2012), *Developing scalable PHP applications using MongoDB*, available at http://www.phpclasses.org/blog/post/118-Developing-scalable-PHP-applications-using-MongoDB.html
- Cristian Bucur (2010), *Folosirea bazelor de date non relationale, NoSQL in aplicatii web*, Simpozionul Internaţional Dezvoltare durabila in conditii de instabilitate economica, Ediţia a II-a, 18 – 19 iunie 2010, supliment al revistei Calitatea – acces la succes, vol. 11, nr.113, 2010, Satu Mare, ISSN 1582-2559
- Cristian Bucur, Tudorica Bogdan (2011), *Solutions for Working with Large Data Volumes in Web Applications*, Conferinţa Internaţională de Informatică Economică - IE 2011 „Education, Research & Business Technologies", **5-7** Mai 2011, Printing House ASE, Bucharest, 2011, ISSN 2247 – 1480
- Emil Eifrem (2011), *NOSQL: scaling to size and scaling to complexity*, available at http://blogs.neotechnology.com/emil/2009/11/nosql-scaling-to-size-and-scaling-to-complexity.hml
- Ernst, Johannes (2010), *SQL Databases v. NoSQL Databases comment*, Communications of the ACM, Volume: 53, Issue: 4, Pages: 11-11
- Kang Wei, Tan Sicong, Xiao Qian, Hadi Amiri (2010), *An Investigation of No-SQL Data Stores*, School of Computing, National University of Singapore
- Vikram Vaswani (2012), *Getting Started with MongoDB and PHP*, available at http://devzone.zend.com /1730/ getting-started-with-mongodb-and-php
- *** (2012), *A PHP MongoDB Admin script for managing databases, collections and documents from a standalone easy to use script*, available at http://github.com/jwage/php-mongodb-admin
- *** (2012), *MongoDB*, available at http://www.mongodb.org/
- *** (2012), *Php-MongoDB driver manual*, available at http://us2.php.net/manual/en/book.mongo.php
- *** (2012), *Scraper class for collecting e-mail addresses*, available at http://www.phpclasses.org/package/5663-PHP-Crawl-pages-and-scrape-e-mail-addresses-into-MySQL.html
- *** (2012), *What is MongoDB?*, available at http://www.10gen.com/what-is-mongodb
- *** (2012), *Z-WAMP Server Pack*, available at http://zwamp.sourceforge.net/

INCENTIVES AND ENHANCEMENTS FOR USING ONLINE TOOLS. CASE STUDY FOR AN ONLINE ASSESSMENT PLATFORM

CERVINSCHI Cezar Liviu
Ph. D. student, Faculty of Cybernetics, Statistics and Economic Informatics, Economic Informatics Department, "Academy of Economic Studies", Bucharest, Romania, cezar@symbolsoftware.ro

BUTUCEA Diana
Ph. D. student, Faculty of Cybernetics, Statistics and Economic Informatics, Economic Informatics Department, "Academy of Economic Studies", Bucharest, Romania, dianabutucea@gmail.com

Abstract: *Nowadays, defining a term like 'incentive' represents a major challenge when the person that needs motivation into doing anything is the online user. In a society where data services have became widely used and where the internet user is an extremely accurate judge, online tools must adapt and keep up with the user's expectations. In this context, the paper will focus on presenting incentives and enhancements technology has brought in order to meet the customer's needs in terms of interactivity, usability and engagement. A case study on online assessment tools will also be presented as part of this research.*

Keywords: online usability, online assessment, usability enhancements, e-learning

JEL classification: C83, C88, C89

1. Introduction

Having to discuss about the two main concepts of this paper – *incentive* and *enhancement* – requires at least two main directions for defining them and the context they will be put in: the social context and the technological context of these terms. Disregarding the purpose of the presented paper, both of these terms can become the subject of several debates over what they truly mean and in which context can they be placed. Analyzing the social aspect of this matter, the term *incentive* can be defined as something that makes a person to perform an action (e.g. someone would go shopping for a certain brand of cell phone because the marketing strategy of the producer is good enough to determine that person to achieve the product). Following the same logic, an *enhancement* is something that adds extra-value to a person or to a product, increasing its value by improving it (e.g. someone would go shopping for a new generation of a cell phone because new features have been added to it and because its overall functions would fully satisfy the potential client). In parallel, the technological facet of the two target terms need a similar analysis, as they stand at the base of the unexpected growth of Information Technology in the past 20 to 30 years. Since people became more and more willing to adopt technology in all areas of life (starting from medicine and ending with home appliances), they needed an incentive that would make them do this faster. This particular incentive came exactly from the area of technology and it offered the end used products that would ease his way of life and the way he would interact to another person (e.g. cell phones, instant messaging , social networks) or to certain machines (e.g. medical equipment, CNCs, home appliances). The growth of technology did this - and it is still actually running this process, but on a different level - it raised this challenge to the consumer, and it was the consumer's choice to adopt it. The incentive has been there all along, it exists in the human nature, but, with all the enhancements technology has brought into day-to-day life, it came natural for the consumer to want more, to expect more and to become extremely selective into satisfying his needs.

Keeping this idea in mind and focusing directly on one of the most important overall component of technology today – the online environment, it can be clearly stated that incentives and enhancements for using it have been constantly added and that new features are developed with high frequencies in order to meet the online consumer's expectations. In the same time, these expectations grew by the day, becoming higher and higher, once technology became more and more sophisticated. Therefore, when talking about the online environment, there are four stages in the evolution of the Internet: pre web, reactive web, interactive web and integrative web (Cervinschi, Butucea, 2010). At this point, the online user is not only able to read the content (static html pages), he can also create content (online web design tools or patterns), add or erase content (e.g. blogs, social networks, forums) and several other

functionalities he can perform, all of them developed and launched in market just to raise to the user's expectations.

The internet penetration worldwide has been massively increased in the last few years. At this point, an impressive list of ways to access data services can be presented (DSL connections, broadband connections, wireless high speed connections, 3G connections, HSUPA/HSDPA connections for mobile data being the most relevant of them), therefore an online user would not only limit his options into accessing the Internet from his PC at work or at home. He would want to have constant access to information on his mobile devices, therefore, in the context of the presented paper, the incentive technology has raised is data access anywhere and at any time.

In this context, what we wanted to achieve in the presented paper was to create a series of profiles for the online users, in relation with the expectations they have from the online environments. This profiling process would find itself extremely useful in our future research when impact assessments over two major software platforms will be done both technologically and socially. At the same time, the paper will present a set of facts that grew the Internet into what it is today. Combining all of these enhancements of data services with the need of higher speed, better interaction with web pages or custom developed frameworks, made the Internet into what it is today and rose it to what the consumer needed. Next, the paper will present a case study over the two main concepts described – *incentive* and *enhancements* – in relation with an online assessment tool developed by the authors, pointing out user feedback as well as pros and cons over the entire software platform. Conclusions and future work will close the purpose of the presented paper.

2. The online consumer

Over the past few years, the Internet evolved in an astonishing rhythm. Due to high expectation the online consumer raised in this period, technologies needed to provide an offer that would satisfy this high demand. And the demand was quite complex, since, from the early beginning and spreading of the Internet, the online user wanted more, being completely unsatisfied with basic static content. If the research done by (Chu et. al, 2007) did a segregation between the stages of online evolution, in terms of user needs and demands, research has found that they kept being higher and higher in most of the cases, online assessment tools being one of them (Paechter et. al, 2010).

Discussing about the online consumer, a few patterns will be defined, based on the author's experience in the research field as well as on customer expectations premises. Starting from the idea that there are three major types of learners (since the paper relates to online assessment) described by specialists and researches (Lecture notes, 1999; LearningRx, 2011), visual learners (see), auditory learners (hear) and kinesthetic learners (touch) and by drawing a parallel with the overall online consumer, the paper will categorize him in three distinct categories, similar to the ones presented – visual consumers, auditory consumers and kinesthetic consumers.

2.1. Visual online consumers

This type of consumer is mainly targeted on what he sees. His best feedback is given when the design or presentation of a webpage raises to his expectations. If the presentation satisfies his needs, than the content falls in the second plan and he would give positive feedback on the overall web content, even though the information does not raise to the presentation. This type of users are not fond of classical, well-structured websites, where relevant information is presented in a clean, simple manner and find these type of websites boring or uninteresting. Their main focus is presentation and design and the usual website pattern they would choose to browse is mainly artistic (arts, photography, landscaping, abstract, non-business layouts).

2.2. Auditory online consumers

This second type of consumer is the consumer that needs to hear something from the website he browses. Even if it is small flash sound when hovering over a button, or live streaming, this user would prefer a website that he can hear and, subconsciously, interact with him. Similar to auditory learners, auditory online users need to hear something in the background while browsing the Internet, and, if this requirement is fulfilled, then they would be able to give positive feedback to what they have browsed on. Their focus is not presentation, but relevant, clean content, pumped-out with small effects that would make the interaction a bit more pleasant. This type of users would prefer Flash, jQuery or AJAX-based templates due to their dynamically generated content (both, consumer or business-oriented layouts).

2.3. Kinestethic online consumers

The third type of online consumer is, at this point, the most widely spread one, because of its characteristics. Since the '*learn by doing*' method can easily be translated to the online consumer, it is pretty obvious that, when performing certain actions and when interacting with a web page, the engagement and the user's will to continue browsing and using the particular page will exponentially increase. This is what the integrative web has brought to the end-user, making him part of the action and a content manager over particular parts of the website. Platforms like Facebook (launched in February 2004 and ranked 2nd in Alexa ranking in April 2012), YouTube (launched in February 2005 and ranked 3rd in Alexa's same ranking) or Twitter (launched in July 2006, ranked 8th) started to engage their users into using them by allowing the consumers several services that include comment threads, file upload, URL sharing, live streaming a.s.o. This type of user does not necessarily follow a certain web design pattern, instead he chooses to provide positive feedback to websites that allow him interaction (e.g. forums, blogs, social networks).

3. Incentives and enhancements into using online tools

Having to meet the expectations of the three user typologies mentioned above has raised several issues for the online industry (but not only for it). On one hand, some major players like Yahoo!, who, at some point, was the indisputable leader of portals and web-based services, suddenly lost his major advance against competition and found himself in a runner-up position sooner than anyone expected. In a similar situation, another technology mammoth, the Finnish mobile phone producer, Nokia Group, is now a runner-up over competition, and being forced to have massive restructuration on both top management and low-level employees, in order to re-align to the industry trends. On the other hands, following the same logic, Google, Facebook, Youtube have massively increased their turnovers because they realised what online users needed – interaction and content. The same thing happened in telecoms industry with American producer Apple, who irreversible changed the mobile phone industry, with the premium high tech gadget that is now a worldwide brand, the Apple iPhone.

Now, coming back to the online industry, all of the three giant corporations mentioned above needed to create an incentive for the user to be engaged and to want to interact with their platforms. This section would cover what technology has brought to the user in order to fulfil his demands and to make the online market a place to use daily.

3.1. Technology growth

As mentioned in the introduction of the paper, technology has brought several benefits in day-to-day life and in all of its domains. Of course, the online domain has not been left aside, since, in the last few years, it had the most rapid growth in overall consumers. In this case, technology needed to adapt to the consumer's demands, therefore new development toolkits have been released into production by giants like Microsoft (Visual Studio), Sun (Java Platform), IBM (Web Sphere Application Server) or Oracle (Weblogic Portal), in order to engage consumers to start using their tools (a free-of-charge trial period) and to agree to buy full enterprise licences (either by using a prepaid method or a pay-as-you-go monthly fee subscription). What these tools have special is that they provide the background for companies to create, produce and manage small (*SOHO* – Small Office/Home Office or *SME* – Small / Medium Enterprise) to huge (*enterprise*) software products.

This approach is, actually, two-sided. On one hand, the customer could receive what he needed in order to start producing software for third party companies, but on another hand, he could, at some point, find himself in a financial collapse, not being able to pay the subscription to the framework provider. For this type of situations, SDK (Software Development Kit) producers incentivize its customers to buy full licences of the product and, if some sort of unavailability to pay arise, they would end their support for the products. This way, the SOHO / SME lose the producer's support but it is still able to take advantage of the development kit he already purchased. In any case, until something bad happens, the customer would always benefit from software and platform upgrades done by the producer, therefore these enhancements would incentivize him into wanting to keep his subscription opened and to receive the latest technology, in order to produce cutting edge products for his own consumers, therefore developing a win-win business relation.

3.2. Software as a Service (SaaS)

The concept behind SaaS has been presented as an improved version of the Application Service Provider (ASP) model (Gonçalves, Ballon, 2011) by the Software Industry Association (SIIA, 2001). Its main benefit is that it offers customers access to different software applications through a network, in a one-to-many architecture. In the last years, it has evolved into an online (web-based) application interface, allowing its customers to easily access, manage and configure the resources they have been given access to, without having any issues in terms of physical location of the machines or applications. This business model has proven itself extremely reliable, since lots of big corporations have adopted it. Since the subscription for SaaS can be either monthly or in pay-as-you-go system, the concept has proven its financial efficiency as well as the functional one, by reducing the customer's hardware resources used in this situation (by using the SaaS platforms and hardware).

However, in terms of associated risks, the SaaS model raises a few flags regarding management of billing and customer data, security, support services, usage meter or service provisioning (Gonçalves, Ballon, 2011). Putting these examples together, they would change into quite an important impact on software reliability or stability, but all of these risks should be minimized and mitigated by performing a set of simple steps in regards of hardware maintenance and resource allocation. By having a good partitioning mechanism, a SaaS provider would always be able to offer his customers reliable services and be sure that his revenue would be constant.

In the paper's context, the incentive of this type of services is represented by a model that allows access to cutting edge technologies or software platforms and all of their benefits, and that customers, by paying the subscription fee, would receive real time support for debugging or production issues. The model applies best for SOHOs and SMEs, but it is has not rarely been seen that even big corporations choose to use this model instead of hosting an in-house dedicated IT department.

3.3. Platform as a Service (PaaS)

The PaaS model represents a new and enhanced version of SaaS, and it offers either the option to build new software or services, incremental, over an existing SaaS platform, either to start developing new applications from scratch. By offering underlying services, PaaS platforms facilitate the entire software lifecycle in terms of application development, design, testing, deployment and hosting (Mitchel, 2008). In terms of application facilities, PaaS allows the runtime environment to run external applications inside the platform itself, meaning that, if the consumer needs it, the platform will be able to access external APIs (Application Programming Interface) or API plug-ins and provided the consumer the expected behaviour of his application. A good example of this approach is the Google concern, that provides developers a set of APIs to access specific zones of their services. Regarding the payment methods and revenue gain, the model charges the customer either on the time spent while using the application engine, either on the time the online user spends browsing the website created through the platform. Also, differentiated quotas can be defined by the service providers, in order to meet the customer's needs.

Using this model acquires a massive advantage over competition – developing and running cloud solutions without having to maintain three different environments (e.g. a development environment, a testing environment and a production one), since all of them would be hosted under the platform's physical machines and the customer would take benefit of them whenever he needs it (Gonçalves, Ballon, 2011). Also, by having a quick time-to-market, the user of such a model might rapidly take the lead in his area of expertise by promoting his particular services and products.

Some of the risks this type of model raises is a slightly difficult portability (if a customer decides to switch from one PaaS provider to another, he might find this procedure a bit unfriendly and he might also be forced to start building his web applications from scratch), thus increasing the development effort. The concept of building once and deploying many opened new possibilities in this case, but it is the PaaS services provider to choose whether he makes his APIs public so that web applications developed on different platforms can be deployed on his or not.

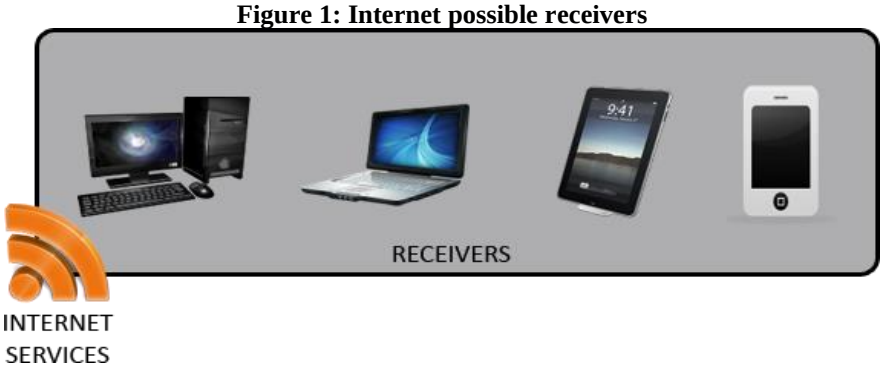
The incentive over using this type of services is represented by the balance between the free and paid features that the provider gives access to but it is, without a doubt, a good alternative over buying full software licenses. In terms of enhancements, this type of model, provided by several players in the market, would only benefit technology, increase and improve the way software systems and web applications are developed and maintained.

3.4. Internet and mobile data services

A discussion about online behavioural models of users or about architectural service platforms would be incomplete if the data services aspect would not be put in the same context. Since the online user, the developer or the software services provider are all bounded by the existence of the world wide web, it becomes almost mandatory to open a topic in regards to this subject. From the early beginning of the internet, people have been fascinated by its features (Butucea, Cervinschi, 2011), and, nowadays, it became an information provisioning tool without which several business branches would definitely fall (e.g. stock markets, e-mail or application servers, web portals a.s.o.).

In this context, the need for online data has increased and, considering that lots of worldwide corporation have offshore offices in other areas of the globe and that all consumers need access to information, communication between data stores, data hubs, cloud servers or any other type of data transmission tools must permanently exist. On a smaller scale, for individual consumers, dial-up, broadband, wireless, 3G, HSUPA/HSDPA services have became commodities and are taken for granted (separately or bundled in packets of services containing voice, television, satellite and data) and the online user has became extremely selective into what he expects to see on a web page and what would be an incentive for him to continue browsing or to come back again on a particular web page (Cervinschi, Butucea, 2011). In order to raise to his demands and expectations, data services providers needed to re-adapt and offer integrated bundles of services that would include broadband internet connection, '*internet on your mobile*' (IOYM), '*internet on any device*'. This means that, at this moment, the only difference between a personal computer or a laptop and a tablet or a smartphone remains the connection speed. Basically, at this point, any mobile device, be it a laptop, a tablet or a mobile phone (smartphone, usually, but not necessarily) can be a good replacement for the old-fashioned PC (Figure 1).

Figure 1: Internet possible receivers



4. Case study

For research purposes, an online assessment tool developed by the authors has been used. What we needed to point out were the strong and weak parts of the software system that would determine the end-user to continue using this particular assessment tool. The feedback has been gathered from a sample of 50 students (aged between 20 and 23) coming from two areas of study – 25 from Computer Science and 25 from Economic Studies. Because their online experience is not that high, meaning that only 20% (10 out of 50) of the students are actually developing online applications, we considered this number of responses relevant and decided to validate the results.

Methodology:

Two days before the examination date, the students have received e-mails with details regarding the address they needed to access over the internet. There were no restrictions over where from to access the specific address, but there has been a limitation for a maximum of 50 opened connections at the time of the exam, set up in the backend of the online assessment software system. Also, the connection channel has been secured and it only ran through https protocol. If, somehow, more than 50 registration requests would hit the application server, then the application would automatically block each request sent besides the 50 that were permitted. Also, no restrictions over what device to use when taking the exams has been raised, therefore students were able to use any type of device, as long as it had an open connection to the internet.

After the successful login of all the students, the administrator of the web application authorized them, un-blocking their first exam, out of the three exams that have been scheduled to be run that day, for experimental purposes. Of course, the disciplines were different for each type of users (Computer Science

or Economic Studies) and the results had no relevance on the students' grades, since the study targeted to measure incentives and enhancements of the online assessment tool that would make the users want to use the same platform again and would choose it instead of another assessment method (e.g. the classical pen and paper evaluation).

Questionnaire:

After closing all three tests, each of them lasting 30 minutes (a total of 90 minutes overall plus the eventual times for each registration to the exam, approximately 5 minutes each, so, in less than two hours – 100 minutes, the students managed to pass through three online exams), the connection with the students has been closed and they had to provide feedback over the assessment methodology by filling up an online questionnaire. The questionnaire contained general questions like "*What is your domain of study?*" or "*What internet browser have you used while taking the online examinations?*" but also critical questions that make the core of the whole paper: "*Would you prefer a classical assessment instead of the online assessment you have just taken?*", "*What did you find attractive or appealing in the assessment tool you have just used?*", "*What new features would you like to see implemented in the assessment tool you have just used?*", "*Have you observed bad or unexpected behaviour of the assessment application during your three tests?*". A full version of the questionnaire is available upon request.

Results:

The results shown that, despite the usability enhancements that have been carefully implemented and treated both at the design level as well as programmatic, some of the subjects found it difficult to browse through the question at the beginning of the first test. Despite that, most part of the students would definitely choose this type of assessment instead of the pen and paper or oral examination methods due to its stringency, clarity and non-time-consuming structure. Also, a main incentive into reusing this online assessment tool, was the fact that the students were not forced to be in the classroom, and they were able to access the platform even by using their mobile phone. This has been the key question of the quiz and the answers we received confirmed what we expected and what we hoped to receive from our research – the confirmation that if a tool has been translated to online, then it would be chosen by the targeted people, the online users, in a world where data flows within mili-seconds and where time is the most important resource on the planet. A full version of the results report is available upon request.

5. Conclusions and future work

The presented paper wanted to provide arguments both theoretical, by exposing a state of the art in means of online environment nowadays, and practical, by describing the case study done using an online assessment tool, on a sample of 50 students that provided feedback over incentives and enhancements they have found while interacting with the web application and that would make them want to re-use the software platform.

The proposed methodology obviously has its advantages and disadvantages, but after running it, we were able to validate the results due to the honesty the sample of students answered our questionnaire, thus being able to also validate our assumptions: if an offline activity is translated to online and if that particular web-based software system has a user-friendly design, some key core functional features that would be able to successfully replace the offline activity, then, at least a part of the population would adopt it – the online users (implicitly, the target of the migration to online).

In conclusion, we can clearly state that our research and main target has been reached and confirmed. Future work in regards to this subject would imply day-to-day research in the fields of online and web technologies and similar studies will be conducted in the field of e-learning software systems. Online cutting edge technology will be strongly monitored and studies regarding web-based training (WBT) will also be made. We also want to approach the meta-data subject, since online assessment tools and e-learning platforms can easily be integrated with this type of segregation mechanism. Nevertheless, a study over a Data Warehouse has already been started and we expect to see results in the next period and to put it into correlation with the existing online assessment tool.

6. Acknowledgements

This work was cofinanced from the European Social Fund through Sectorial Operational Programme Human Resources Development 2007-2013, project number POSDRU/107/1.5/S/77213 „Ph.D. for a career in interdisciplinary economic research at the European standards".

7. References

- Butucea, D., Cervinschi, C. L., Building Robust E-learning Software Systems Using Web Technologies, *Informatica Economica*, 15, 2011, pp. 1107 – 118
- Cervinschi, C. L..; Butucea, D. (2010) Integration of Web Technologies in Web 2.0 Applications. Is Web 2.0 a Solution?, *Database Systems Journal*, 2, 2012, pp. 39 – 44
- Cervinschi, C. L., Butucea, D., Impact of web technologies in functionality and aesthetics of web applications, Revista Economica, 5, 2011, pp. 12 – 17
- Chu, S. C.; Leung, L. C,.; Hui, Y. V.; Cheung, W., Evolution of e-commerce Web sites: A conceptual framework and a longitudinal study, *Information & Management*, 44, 2007, pp. 154–164
- Gonçalves, V.; Ballon, P., Adding value to the network: Mobile operators' experiments with Software-as-a-Service and Platform-as-a-Service models, *Telematics and Informatics*, 28, 2011, pp. 12 – 21
- LearningRx LLC, 2011 LearningRx Results Report, LearningRx LLC 2012, available at http://news.learningrx.com/results/, accessed on the 20th of March, 2012
- Lecture Notes - Week 10, Learning - Chapter 14, TYPES OF LEARNING, available on http://home.epix.net/~tcannon1/Physioweek10.htm, accessed on the 20th of March, 2012
- Mitchell, D.,Defining Platform-as-a-Service, or PaaS, 2008, available on http://blogs.bungeeconnect.com/2008/02/18/defining-platform-as-a-service-or-paas, accessed on the 22nd of March 2012
- Paechter , M.; Maier, B.; Macher, D., Students' expectations of, and experiences in e-learning: Their relation to learning achievements and course satisfaction, *Computers & Education*, 54, 2010, pp. 222 – 229
- SIIA, Software as a Service: Strategic Backgrounder, Software & Information Industry Association, Washington, 2001
- http://www.wikipedia.org/, accessed on the 20th of March, 2012

ON A MODEL FOR PREDICTING THE EXCHANGE RATE EURO-LEU WITH A NAR NEURAL NETWORK

CIOBANU Dumitru
Ph.D. student, Faculty of Economics and Business Administration, University of Craiova, Craiova, Romania, ciobanubebedumitru@yahoo.com

Abstract: *Developing new methods for predictive modeling of time series and application of existing techniques in many other areas will be a permanent concern for both researchers and companies that are interested to gain competitive advantages. In this paper, I used Matlab software suite to create a NAR (nonlinear autoregressive) neural network able to predict the following values of the series of exchange rate euro-leu. Using graphs obtained from numerous simulations emphasize the chaotic nature of the series. I also explore the possibility of improving the predictions with a time series made of arithmetic averages of predictions.*

Key words: exchange rate, neural network, Matlab, chaos.

JEL classification: C45, C53, C63.

1. Introduction

Most companies and organizations from today collect data on a large scale. This huge amount of data can be used to extract knowledge that can represent a real advantage to business managers. Analyzing accurately and timely of this large amounts of data is a difficult task, in general, not possible with traditional methods. Ability to analyze and use the massive amounts of data remained far behind the possibilities of storing them. This raises new challenges for businessmen and researchers for the purposes of extracting useful information (Wang, Liao & Rees, 2002).

Increasing accuracy for predictions can save significant amount of money for a company and is a major motivation for using the methods of forecasting and systematic investigation for new models and techniques to improve the current results (Zhan, 2003).

If a number of explanatory variables must be identified and predicted, using time series approach has the advantage of easy preparation of data collection and modeling.

In time series prediction, historical data is collected and analyzed to produce a model that capture the relationships between observed variables. The model is then used to predict future values of time series. There have been numerous efforts to develop and improve methods for time series prediction.

Linear approach assumes a linear process that characterizes the data generation. There are a number of linear prediction models such as moving average, exponential smoothing, time series regression and time series decomposition.

One of the most important models is Autoregressive Integrated Moving Average - ARIMA, which was developed by Box and Jenkins (Box & Jenkins, 1976) in the 70s. Often, ARIMA is called Box-Jenkins model.

Although ARIMA models are quite flexible and can be used for a large number of time series the main limitation is given by the assumption of a linear form for the model. This means that an autocorrelated linear structure is supposed to be granted before according the model with data. Therefore, an ARIMA model cannot capture nonlinear behavior that is quite common in economic time series. Approximation with linear models of complex real-world problems is not always satisfactory as was stressed at a publicized M-competition in the early 80 (Makridakis, et al., 1982).

The approach of nonlinear modeling the time series is suitable for most real-world problems. The real world is nonlinear and complex rather than linear because there are so many possible nonlinear structures and relationships. Most nonlinear models developed during the last two decades are parametric. In order to use the parametric model, it must be specified before. Therefore, they cannot be used if the data characteristics do not match the model assumptions. Parametric approach is very suitable for nonlinear problems with a complex structure but there is a lack of theories that suggest a specific form for the structure.

Artificial neural networks are algorithms and techniques that can be used for statistical modeling and is an alternative to linear regression models, the most common approach for developing predictive models.

Neural networks have several advantages including less need for formal statistical training, ability to detect, implicitly, complex nonlinear relationships between dependent and independent variables, ability to detect any possible interactions between predictor variables and the existence of a wide variety of training algorithms.

Disadvantages of neural network include the nature of "black box" computing, inclination for memorize the data (network loses the ability to generalize), and the empirical nature of the model developed.

In this paper, I used Matlab programming environment to create and train a neural network capable to predict the following values of the exchange rate euro-leu series. I used the built network to determine the following predictions of exchange rate and analyzed the evolution of predictions compared with observed data.

2. Neural Networks

Artificial Neural Networks (ANN) are information processing systems, composed of simple processing units, interconnected and operating in parallel. As in nature, network function is determined by the connections between elements. Weights of links between units are those that store network information. The network is trained by adjusting these weights, according to an algorithm.

An artificial neural network, or simply neural network (NN), is a set of information processing elements inspired by the functioning of biological nervous systems, like the processing of information by human brain. The key element of this field is the new structure of information processing system. It consists of a large number of processing elements (neurons) working together to solve a specific problem.

Characteristics of Artificial Neural Networks (ANN):

- **Distributed representation of information**: the information stored by the network is distributed in the structure of weights, which makes that the effect of certain inputs on output depend on all the network weights.

- **Ability to generalize**: for situations that was not found in the training data. This feature depends on the number of weights, i.e. the network size. It is found that increasing network size leads to a good training data memory, but decreases performance on test data, which means that ANN has lost the ability to generalize. Determining the optimal number of hidden layer neurons, which is a key step in designing an ANN can be done by choosing the number of neurons from where begins to decrease the performance of ANN on the test set.

- **Tolerance to noise**: ANN can be trained, even if the data are affected by noise, decreasing obviously her performance.

- **Resistance to partial destruction**: because distributed representation of information, ANN can operate even after the destruction of a small part.

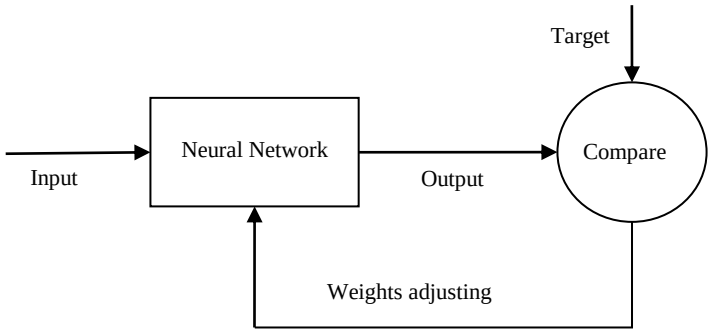
- **Speed into account**: ANN is training time consuming, but once trained will quickly calculate the network output for a given input.

We can train an ANN to perform a particular function by adjusting the connection values (weights) between elements.

ANN is usually adjusted (trained), so that a signal input to involve a certain output (target). Such a situation is shown in

Fig. 0.1. The network is adjusted based on comparison between the response and target, until the network output matches the target. To train a network in supervised training, use several pairs input / target.

Fig. 0.1. Supervised training of a neural network.



Neural networks are universal approximators and works best when used for modeling systems that have high tolerance to errors. However, they work well for:
• capturing associations or discovering regularities within a set of models;
• if the volume, number of variables or data diversity is very high;
• relationships between variables are vaguely understood;
• relationships are hard to describe adequately with conventional approaches.

The strength of neural networks is their ability to accurately predict outcomes for complex problems. In tests of accuracy, compared with other approaches, neural networks are always able to get very good scores (Berson, Smith & Thearling, 1999).

3. Predicting the euro-leu exchange rate

Time series used is the euro-leu exchange rate and was downloaded from the website of National Bank of Romania at http://www.bnr.ro/Raport-statistic-606.aspx.

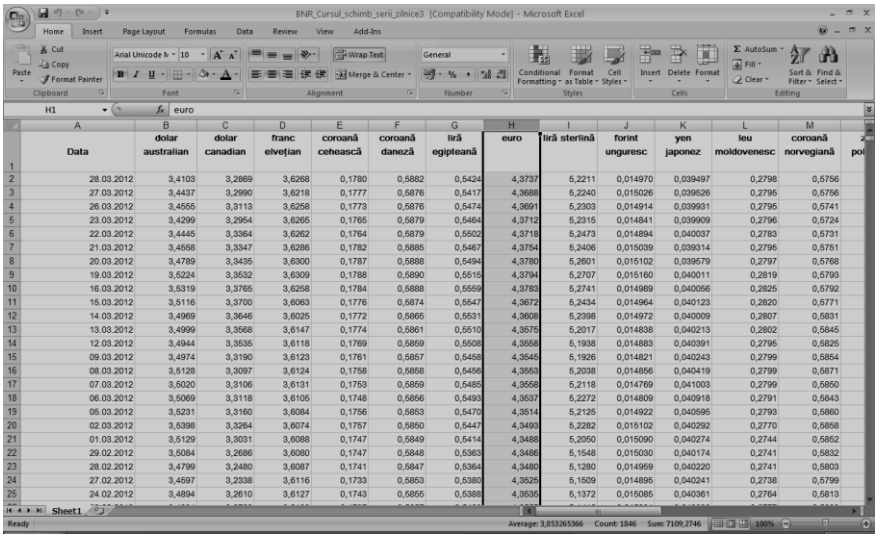
After removal of lines, unhelpful to our analysis, the excel file with values of exchange rate between several currencies and leu is shown in **Fig. 0.2**.

We will focus on the euro-leu exchange rate set on the column 8, the one highlighted in the excel file shown in the screenshot below.

The program used to perform simulations was Matlab version 7.12.0 (R2011a).

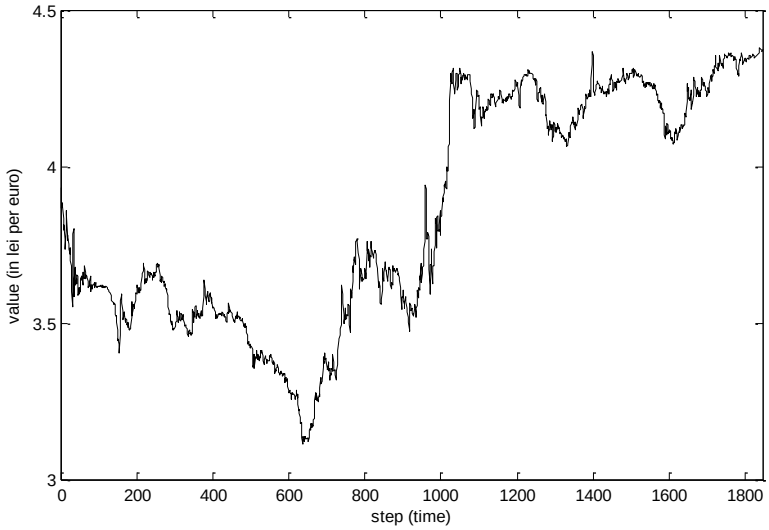
I imported into Matlab data from column 8 of the excel file and I ordered ascending by date, which is exactly the opposite order than in the original file.

Fig. 0.2. A part of the file with the exchange rates between some foreign currencies and leu.



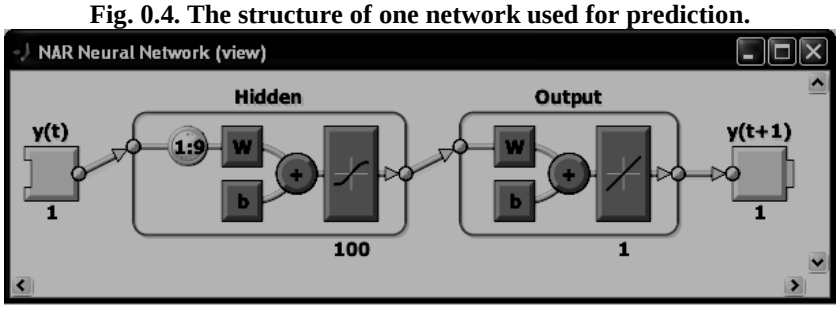
	Data	dolar australian	dolar canadian	franc elveţian	coroană cehească	coroană daneză	liră egipteană	euro	liră sterlină	forint unguresc	yen japonez	leu moldovenesc	coroană norvegiană
2	28.03.2012	3,4103	3,2869	3,6268	0,1780	0,5882	0,5424	4,3737	5,2211	0,014970	0,039497	0,2798	0,5756
3	27.03.2012	3,4437	3,2990	3,6218	0,1777	0,5876	0,5417	4,3688	5,2240	0,015026	0,039526	0,2795	0,5756
4	26.03.2012	3,4555	3,3113	3,6258	0,1773	0,5876	0,5474	4,3691	5,2303	0,014914	0,039931	0,2795	0,5741
5	23.03.2012	3,4299	3,2954	3,6265	0,1765	0,5879	0,5464	4,3712	5,2315	0,014841	0,039909	0,2796	0,5724
6	22.03.2012	3,4445	3,3364	3,6262	0,1764	0,5879	0,5502	4,3718	5,2473	0,014894	0,040037	0,2783	0,5731
7	21.03.2012	3,4568	3,3347	3,6286	0,1782	0,5885	0,5467	4,3754	5,2406	0,015039	0,039314	0,2795	0,5751
8	20.03.2012	3,4789	3,3435	3,6300	0,1787	0,5888	0,5494	4,3780	5,2601	0,015102	0,039579	0,2797	0,5768
9	19.03.2012	3,5224	3,3532	3,6309	0,1788	0,5890	0,5515	4,3794	5,2707	0,015160	0,040011	0,2819	0,5793
10	16.03.2012	3,5319	3,3765	3,6258	0,1784	0,5888	0,5559	4,3783	5,2741	0,014989	0,040056	0,2825	0,5792
11	15.03.2012	3,5116	3,3700	3,6063	0,1776	0,5874	0,5547	4,3672	5,2434	0,014964	0,040123	0,2820	0,5771
12	14.03.2012	3,4969	3,3646	3,6025	0,1772	0,5865	0,5531	4,3608	5,2398	0,014972	0,040009	0,2807	0,5831
13	13.03.2012	3,4999	3,3568	3,6147	0,1774	0,5861	0,5510	4,3575	5,2017	0,014838	0,040213	0,2802	0,5845
14	12.03.2012	3,4944	3,3535	3,6118	0,1769	0,5859	0,5508	4,3568	5,1938	0,014883	0,040391	0,2795	0,5825
15	09.03.2012	3,4974	3,3190	3,6123	0,1761	0,5857	0,5458	4,3545	5,1926	0,014821	0,040243	0,2799	0,5854
16	08.03.2012	3,5128	3,3097	3,6124	0,1758	0,5858	0,5456	4,3553	5,2038	0,014866	0,040419	0,2799	0,5871
17	07.03.2012	3,5020	3,3106	3,6131	0,1753	0,5859	0,5485	4,3558	5,2118	0,014769	0,041003	0,2799	0,5850
18	06.03.2012	3,5069	3,3118	3,6105	0,1748	0,5856	0,5493	4,3537	5,2272	0,014809	0,040918	0,2791	0,5843
19	05.03.2012	3,5231	3,3160	3,6084	0,1756	0,5853	0,5470	4,3514	5,2125	0,014922	0,040595	0,2793	0,5860
20	02.03.2012	3,5398	3,3264	3,6074	0,1757	0,5850	0,5447	4,3493	5,2282	0,015102	0,040292	0,2770	0,5858
21	01.03.2012	3,5129	3,3031	3,6088	0,1747	0,5849	0,5414	4,3488	5,2050	0,015090	0,040274	0,2744	0,5852
22	29.02.2012	3,5084	3,2686	3,6080	0,1747	0,5848	0,5363	4,3486	5,1548	0,015030	0,040174	0,2741	0,5832
23	28.02.2012	3,4799	3,2480	3,6087	0,1741	0,5847	0,5364	4,3480	5,1280	0,014959	0,040220	0,2741	0,5803
24	27.02.2012	3,4597	3,2338	3,6116	0,1733	0,5853	0,5380	4,3525	5,1509	0,014895	0,040241	0,2738	0,5799
25	24.02.2012	3,4894	3,2610	3,6127	0,1743	0,5855	0,5388	4,3535	5,1372	0,015085	0,040361	0,2764	0,5813

Fig. 0.3. The evolution of exchange rate euro-leu, 1845 observations.



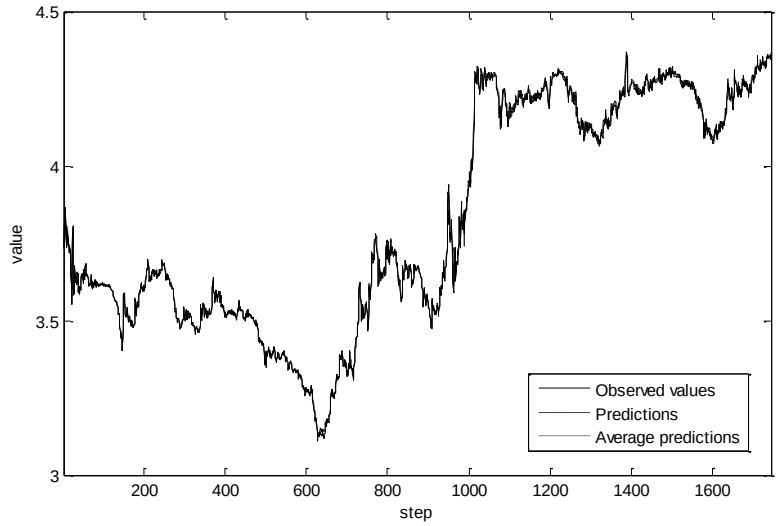
In the figure above is shown the evolution of exchange rate between euro and leu for 1845 observations. From this, I kept 100 observations for tests and use the other 1745 for training the networks.

Fig. 0.4. The structure of one network used for prediction.



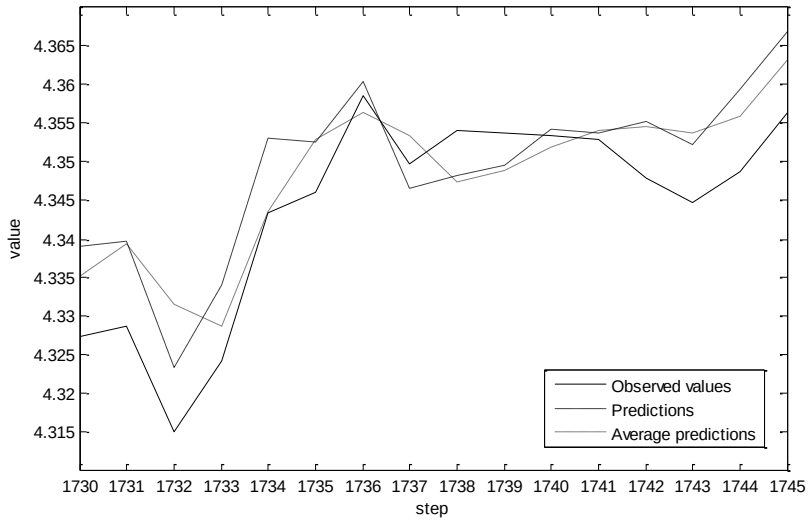
I trained a NAR (nonlinear autoregressive) Neural Network that uses nine past values of the exchange rate and predict the next value of the time series. The structure of the network is presented in **Fig. 0.4**. and consist in 100 hidden neurons with sigmoidal activation function and an output neuron with linear activation function.

Fig. 0.5. Representation of time series and predictions for training set.



In Fig. 0.5. we can see that the predictions are very close to the original time series. To see how it really looks like we must take a closer look, see Fig. 0.6.

Fig. 0.6. A closer look to the final part of prediction.



I measured the network performance using MSE (Mean Squared Error). For the training data considered in my example I obtained MSE = $2.6883 \cdot 10^{-4}$. MSE for the mean predictions was $1.1108 \cdot 10^{-4}$.

I used the neural network to predict the next 100 values. First I used as inputs the observed values, introducing nine measured values, the neural network provides our next prediction for the exchange rate. The results were satisfactory, see **Fig. 0.7** and

Fig. 0.8, with MSE = $5.2317 \cdot 10^{-5}$ for predictions and $5.0690 \cdot 10^{-5}$ for mean predictions.

This was not the only case in which the mean predictions have returned an mean squared error better than the predictions.

Fig. 0.7. Next 100 predictions using as inputs measured values.

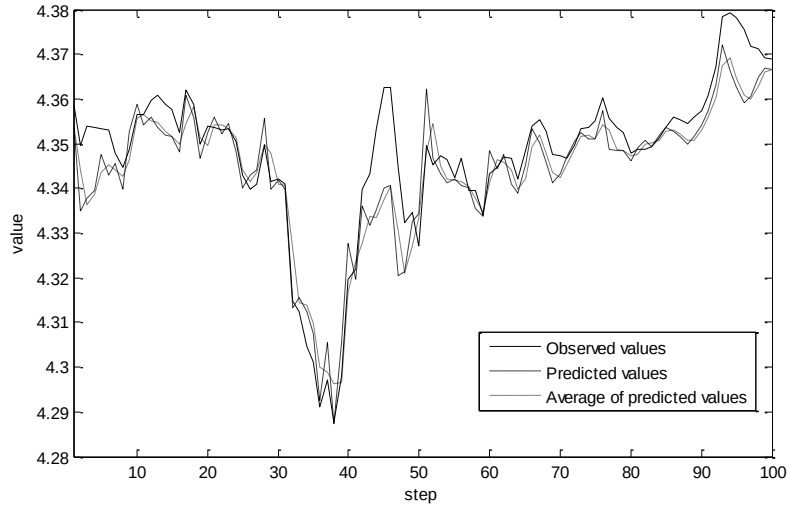
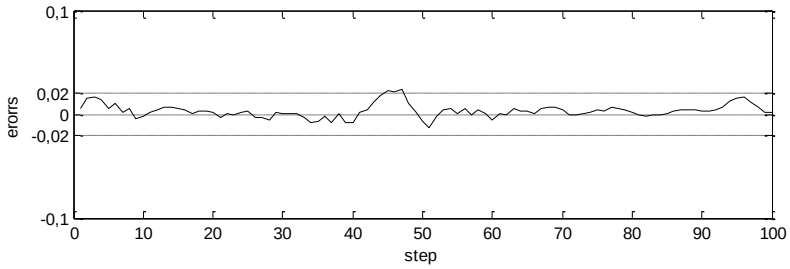
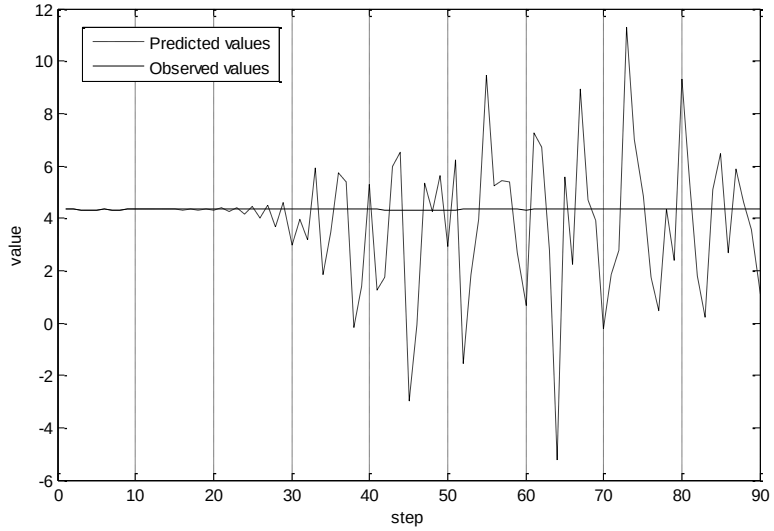


Fig. 0.8. The errors for next 100 predictions.



Then, I used the neural network to predict the next 100 values but using as input the predictions obtained, the result can be seen in **Fig. 0.9**.

Fig. 0.9. Next 100 predictions using as inputs predicted values.



I have performed hundreds of tests and each time the series of predictions showed large deviations from the observed values after a relatively small number of steps. I concluded that all models obtained using neural networks to predict exchange rate shows a chaotic behavior.

Thus, considering the chaotic nature of exchange rate time series, prediction with an acceptable error can be done only a few steps forward.

4. Conclusions

Using neural networks to predict the exchange rate is a good alternative to traditional predictive methods. The fact that predictions for a longer period not working is not a minus of neural networks over other methods but underlines the chaotic nature of time series euro-leu. Due to the chaotic nature of exchange rate time series, prediction for several steps, theoretically, would be possible only with a very complicated model. There is not such a model yet soo other methods are also unable to make such predictions.

In this paper I used only time series of the exchange rate for prediction. To improve predictions, in future works, I will use neural network NARX (Nonlinear AutoRegressive with eXogenous input) and will investigate what factors influence foreign exchange rates. As time series that may influence the euro-leu exchange rate I think to use: euro-dollar exchange rate, the external debt, level of the exports, level of the imports, etc.

It also remains to be investigated in what cases the mean predictions represent an improvement of predictions.

5. References

- Box, G. E. P.; Jenkins, G. (1976) *Time Series Analysis: Forecasting and Control*, Holden-Day, 1976.
- Ciobanu, D. (2011 a) *Software used for Data Mining*, Proceedings of the international conference ECOTREND 2011 – Exit from the crisis and revival of sustainable growth, November 25-26, Târgu Jiu, Romania, 2011.
- Ciobanu, D. (2011 b) *Using Neural Networks for Prediction in Business Planning*, Proceedings of the international conference ECOTREND 2011 – Exit from the crisis and revival of sustainable growth, November 25-26, Târgu Jiu, Romania, 2011.
- Kadilar, C.; Simsek, M.; Aladag, C. H. (2009) *Forecasting The Exchange Rate Series With Ann: The Case Of Turkey*, Istanbul University Econometrics and Statistics Journal, vol. 9, issue 1, pp. 17-29, 2009.
- Makridakis, S.; Anderson, A.; Carbone, R.; Fildes, R.; Hibdon, M.; Lewandowski, R.; Newton, J.; Parzen, E.; Winkler, R. (1982) *The accuracy of extrapolation (time series) methods*, Results of a forecasting competition. *Journal of Forecasting*, 1, pp. 111-153, 1982.
- Sumathi, S.; Surekha, P. (2010) *Computational Intelligence Paradigms: Theory and Applications using Matlab*, CRC Press, ISBN 978-4398-0902-0, New York, 2010.
- Wang, M.; Rees, S.J.; Liao, S.Y. (2002) *Building an online purchasing behavior analytical system with neural network*, Edited by Zanasi, Brebbia and Melli, DataMining III., WIT Press, 2002.
- Widrow, B.; Lehr, M.A. (1990) *30 Years of Neural Networks: Perceptron, Madaline and Backpropagation*, Proceedings of the IEEE, 78:1415–1442, 1990.
- Wilshaw, D.J.; Von Der Malsburg, C. (1976) *How Patterned Neural Connections can be Set up by Self-Organization*, in Proceedings of the Royal Society of London B, vol. 194, pp. 431-445, 1976.
- Yang, N. (2011) *Research on prediction of chaotic exchange rate time series applying dynamic component predicting model*, 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), Zhengzhou, China, 2011.
- Zhang, G. P. (2003) *Time series forecasting using a hybrid ARIMA and neural network model*, Neurocomputing, vol. 50, pp. 159-175, 2003.
- Zhang, G.; Patuwo, B. E.; Hu, M. Y. (1998) *Forecasting with artificial neural networks: The state of the art*, International Journal of Forecasting, 14, pp. 35-62, 1998.

WHY CAN'T WE USE THE SAME CURE FOR ALL CRISES? -AN AGENT BASED MODELLING APPROACH-

CIUTACU Ileana
Ph. D. Student, Economics Department, Bucharest Academy of Economic Studies, Bucharest, Romania, ileana.ciutacu@gmail.com

VÎNTU Denis
Ph. D. Student, Economics Department, Bucharest Academy of Economic Studies, Bucharest, Romania, d.vintu@reading.ac.uk

SĂVULESCU Iulian
Ph. D Student, Accounting Department, Bucharest Academy of Economic Studies, Bucharest, Romania, savulescuiulian@yahoo.com

Abstract: *Albert Einstein once said "everything is relative" and he was right . The results of this paper prove him right even when talking about economics, because in real life when we find ourselves in front of a crisis situation we try to solve or cure it using 'old remedies' or what we know, but what we never seem to learn is that using the same inputs will not give us every time the same results. To reach this conclusion we designed and created a simple economy agent based model that was after that developed using NetLogo.*

Key words: economy, crisis, netlogo, agent based models,

JEL classification: A 10, C 60, E 20, J 00, P 40.

1. Introduction

Albert Einstein once said "everything is relative and he was right, not only in what regards Physics, time and space, but also in what regards most Social Sciences, including Economics. In this last few decades, and especially in the last five-six years due to the subprime crisis in the United States of America that lead to the current mondial crisis, Economics and economies have changed considerably, leaving economic policies decidents in great doubt.

According to Farmer and Foley (2009), these problems occur because most governments don't use sophisticated quantitative computer models to guide the country's economy out of crisis. Instead, they use two main cathegories of models, the first type are econometric: empirical statistical models that forecast what will happen in the future with data from the past, assuming that things will stay more or less the same. And the second type of models are based on the "dynamic stochastic general equilibrium" and imply a perfect world that by their very nature rule out economic crisis. Unfortunately, both have many flaws, mainly because there is no or little attempt to assemble the pieces of an economic system and understand it's behavior as a whole.

As a solution to this problem, Farmer and Foley present a better way: agent based modeling. This is a computerized simulation of different types of agents (from consumers to policy-makers and countries) which interact using prescribed rules in a 'world', an artificial economy, that can have a structure as diverse and complete as is needed. Instead of expecting the movement of the economy towards a predetermined equilibrium state, in the case of agent based models, agents behave all their life according to given rules, what they have learned from past experiences and the situation they find themselves in that moment. Thanks to the computer, this type of models can keep track ot all the agent interactions, so the newly created model of artificial economy can be seen clearly, at any given time, from the smallest detail to it as a whole complex system. Also, agent based simulations are able to handle nonlinear behaviour in a wider range that conventional equilibrium models can, thus allowing policy-makers to simulate and explore the quantitative consequences of different policy scenarios upon an artificial economy.

Agent based models are already used with succes by policy-makers in some areas of science as are epidemiology, traffic control or demographics. There also exist some succesfull models that simulate small portions of the economy, as are the financial market model of Blake LeBaron, the firm dynamics

model of Rob Axtell, the credit sector model of Mauro Gallegati's group, presented by Farmer and Foley (2009) or the EURACE project that models the European economy (Langnick 2011).

Another agent-based model that simulates very well the economy the one built by Langnick and presented in one of his papers (Langnick, 2011). This model is a baseline macroeconomic one, that even if simple, can reproduce a lot of the stylized facts of business cycles and can be an adequate response to recent criticism of macroeconomic methodology. This is because, according to Langnick (2011), his model does not depend on the strict assumption of rationality, allows aggregate behavior and absolutely consistent micro foundations and does not depend on equilibrium assumptions that might rule out by definition coordonation failures, instability and crisis. His model is a stylized one that abstracts from the real economy because it contains only a small number of different agent types and interaction rules and can be run on an ordinary desktop PC. Being built upon the model developed by Gaffeo et all. (2008), Langnick's model has only two types of agents: households and firms which use simple behavioral rules in order to choose prices and wages; the economy is allowed to self-organize toward a spontaneous order and a part of the agents are allowed to buy only from a subset of agents. But, instead of analyzing growth as a result of investment in Research and Development, using another type of agents (banks) and indexing time by quarters, Langnick's model measures time in days and months and aims to define a model that is concerned with basic macroeconomic relations in a non-growth environment.

Agent-based models can be designed and created by using special computer programs. One of these programs is the one built in 1999 by Wilensky and developed over time at the Center for Connected Learning and Computer-Based Modeling of the Northwestern University, named NetLogo. This computer program belongs to the next generation of multi-agent modeling languages like StarLogo or StarLogoT, and is run as a standalone application on the Java virtual machine of computers that use Mackintosh, Windows, Linux, etc. platforms.

When using NetLogo, the creator of the agent based model uses a special programing language that allows him to define an environment for simulating natural and social phenomena by using four type of entities: turtles (mobile agents), patches (static agents), links and the observer. These agents can be part of populations that can number up to thousands, operate independently in an artificial 2D or 3D world using predefined rules and instructions. This caractheristic of NetLogo makes it possible to explore the connection between the micro-level behavior of individuals and the macro-level patterns that emerge from their interactions.

2. The Model

By creating a simple economy agent-based model, this paper aims to see the causes for its entering a crisis. Our simple economy model was build based on the NetLogo Wolf Sheep Predation Model (Wilensky, 1997). In this model, Wilensky created a world in which live three types of agents: one type of patches: the patches of grass, and two of turtles: the sheep and the wolf, therfore two prey-predator systems that interact one with the other. The sheep eat the green grass when they encounter it and gain energy, reproduce according to a predetermined probability and are eaten by wolves. The wolves eat the sheep when they encounter them and gain energy, and also reproduce according to a predetermined probability. The grass is eaten by the sheep and grows back according to the time of regrowth set by the model's creator or user.

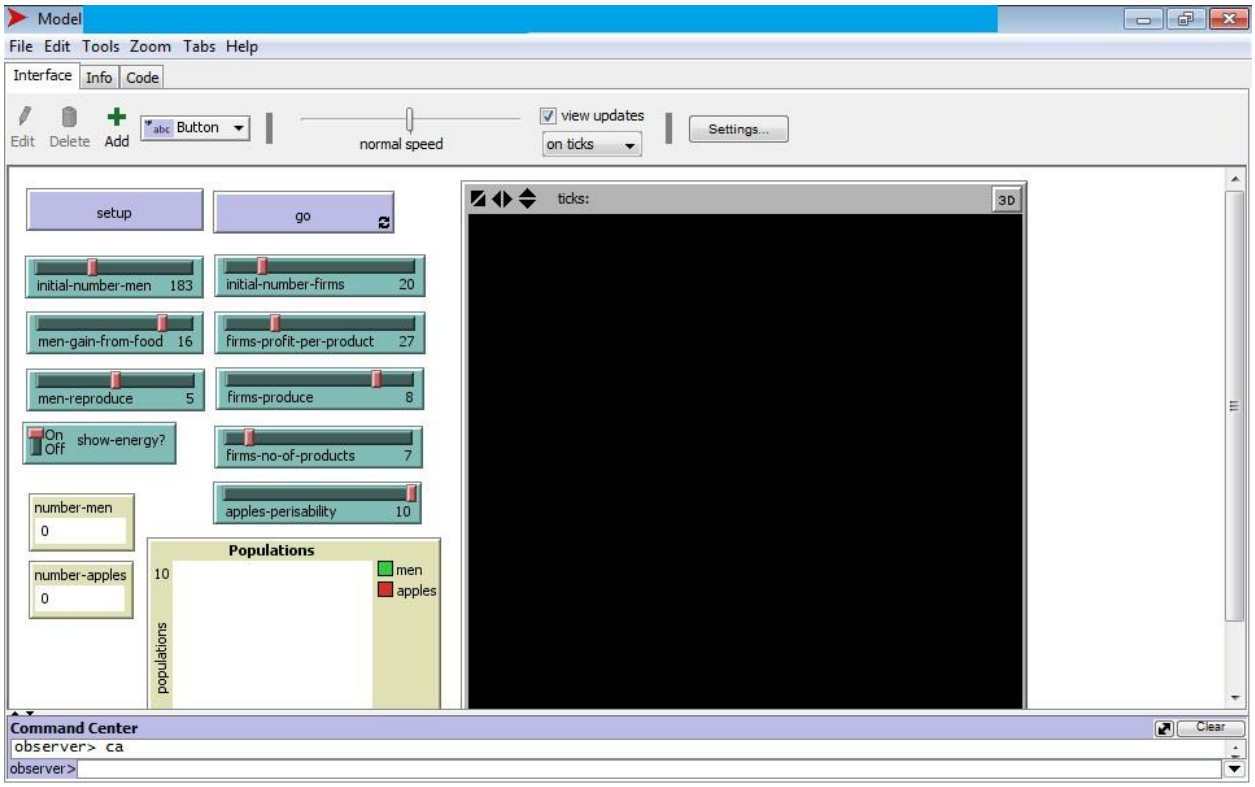
Instead of using two types of turtles and one of patches, the current model has three types or *breeds* of turtles: firms, men and apples, that live on the surface of a sphere-like world. Their position (firms and men) in the world is randomly chosen by the computer every time the model's used presses the 'setup' button.

1. Programming code for the populations

```
breed [men man]
breed [firms firm]
breed [apples apple]
```

The world is a 61*61 units torus, which means that this world does not have borders. Our world's time is measured in ticks by a tick counter (see Figure 1.) and lasts as long as men live in it. When there are no more men, the tick counter stops showing the length in time of that world. The firms and men appear first on the surface of the world in populations of 0 to 100 firms and 0 to 500 men. The currency of this world is energy and can be measured by rational numbers.

Figure 1: The world in the NetLogo interface before pushing the "setup" button



Source: Author's processing

2. Programming code for the economy's 'currency'

```
turtles-own [energy]
```

The firms are born with an initial energy of 250 units, receive 20 units of energy when hiring a man and get their energy cut in half if the odds, determined by a probability and a rule, tell them that they can produce apples.

3. Programming code for the production of goods (apples)

```
to produce-goods
   ask firms [if random-float 10 > firms-produce and energy >= 75 [ set energy (energy / 2)
   hatch-apples firms-no-of-products [ …]
end
```

The apples are linked to the firms, because, even if turtles, apples can not reproduce themselves, but are produced by firms and can live in the world only for 0 to 10 ticks (see the apples-perisability slider in Figure 1.). As in the case of the firms, the apples' position is fixed during the running of a simulation of our simple economy model (or world's time) and is randomly chosen by the computer when the firms that produce them are allowed by the odds and have enough energy for producing apples.

4. Programming code for the "setup" button

```
to setup
   clear-all
   setup-world
end
```

Even if NetLogo turtles have the ablity to move in the world, in this simple economy only the men can change their position during a simulation. By moving around, one step at a tick, men receive energy by getting a one-time tick job, when encountering a firm, or by consuming an apple, when

encountering one. The job cand bring to a man a fix wage of 15 energy units, while consuming an apple can bring the man up to 20 energy units, depending on what the user chooses the men-gain-from-food slider to be before pushing the setup button (Figure 1.). Of course, men can also reproduce, and, just like in the case of firms producing apples, this happens only if they have enough energy and if the odds allow them.

5. Programing code for the men's reproduction process

```
to reproduce  ;; if a man has more than 50 energy and the right probability, it can reproduce
  ask men [
    if random-float 10 > men-reproduce and energy > 50 [
      set energy (energy / 2)
end
```

Men die or get out of our simple economy model's market if they have no more energy left. In that moment, the "go" button, that allows us to run the simulations, stops.

6. Programing code for the men's 'death' of leaving the simple economy's market

```
to death
  ask men [
    if energy < 0
    [die]  ]
end
```

7. Programming code for the go button

```
to go
  if not any? men [stop]
  move
  work
  .
  .
  .
  tick
end
```

In order to see if our simple economy model can get into a crisis, we will set 3 scenarios in which only some variables will change. These variables can be easily changed by modifying the sliders on the interface screen of the NetLogo application (see Figure 1.). We will run each sceario 3 times, until the go button stops, to see what happens and why this happens:

a) First scenario:

- initial-number-men: 150
- initial-number-firms: 10
- firms-no-of-products: 4
- men-reproduce: 4
- firms-produce: 4
- apples-perisability: 5
- men-gain-from-food: 15

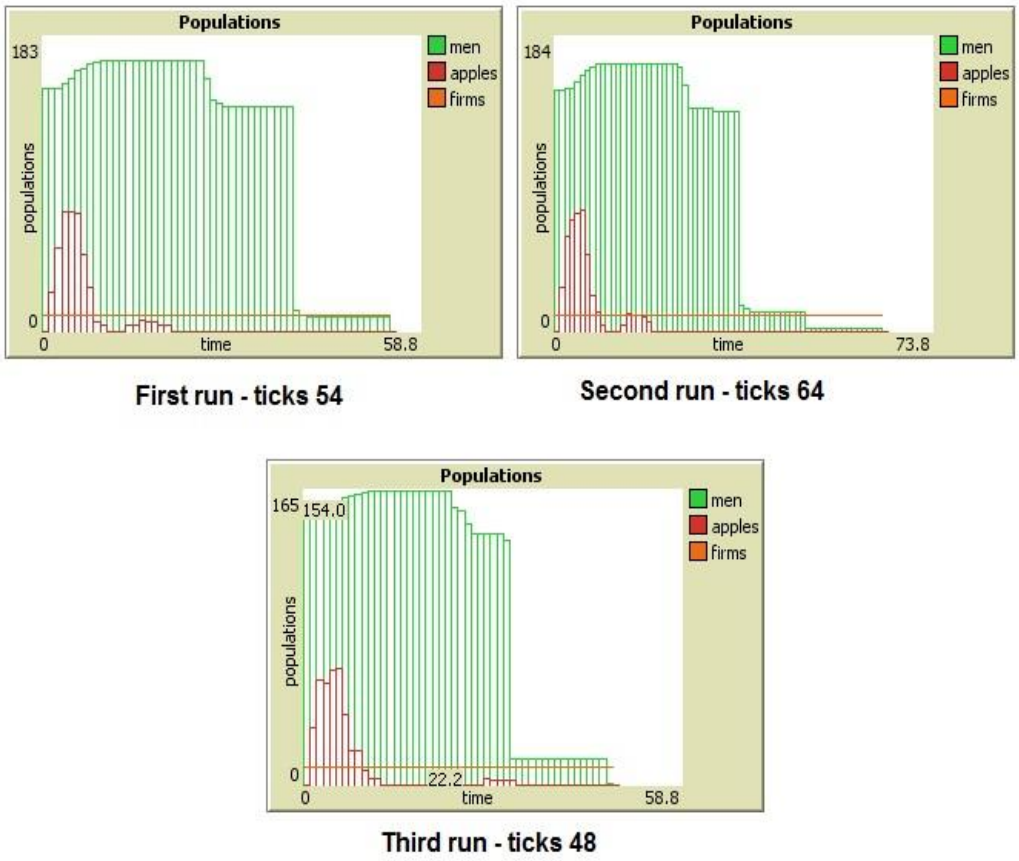
b) Second scenario:

a) initial-number-men: 250
b) initial-number-firms: 20
c) firms-no-of-products: 10
d) men-reproduce: 5
e) firms-produce: 5
f) apples-perisability: 3
g) men-gain-from-food: 15

c) Third scenario:

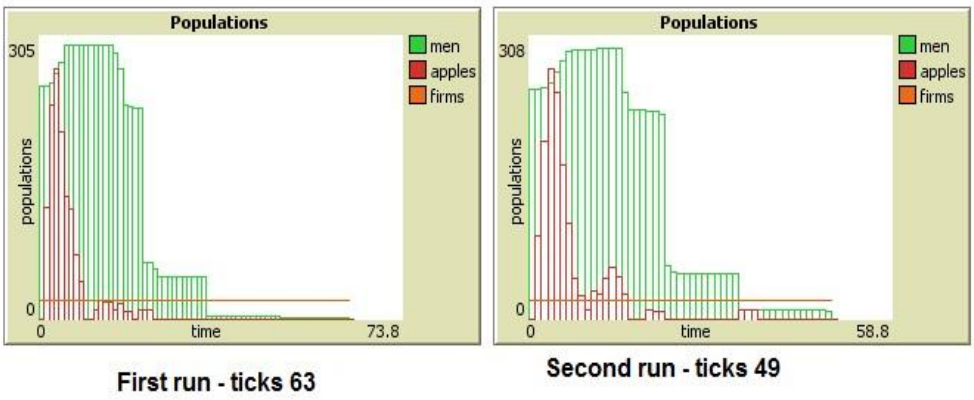
h) initial-number-men: 450
i) initial-number-firms: 40
j) firms-no-of-products: 20
k) men-reproduce: 7
l) firms-produce: 2
m) apples-perisability: 2
n) men-gain-from-food: 10

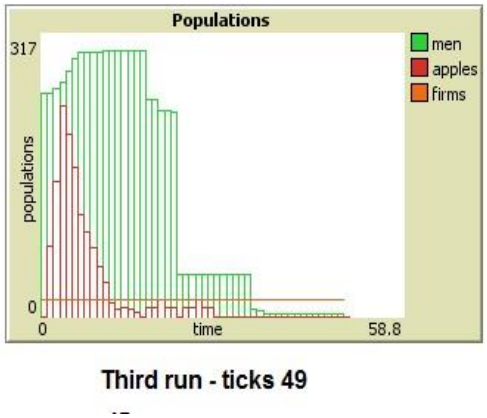
Figure 2: First scenario results



Source: Author's processing

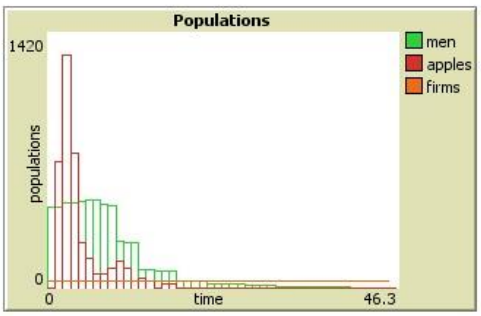
Figure 3: Second scenario results



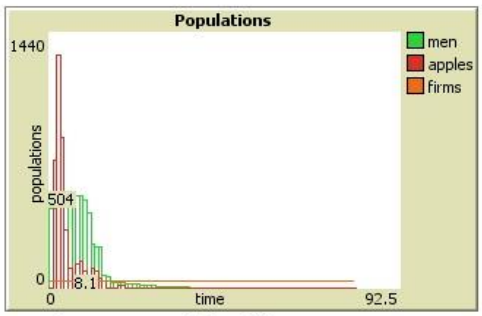


Source: Author's processing

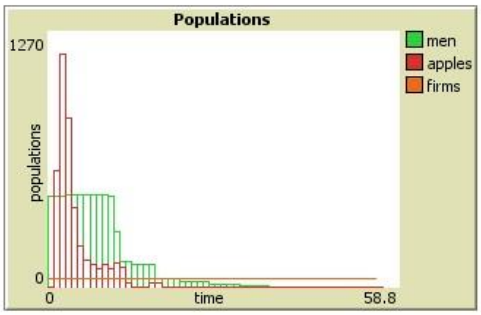
Figure 4: Third scenario results



First run - ticks 45



Secon run - ticks 80



Third run - ticks 55

Source: Author's processing

3. Conclusions

As it can be seen in Figure 2, the three runs of our simple economy model's first scenario, even if have the same inputs, give different results of time and population numbers. Thus, the first run lasted 54 ticks or 58.8 secconds and had a maximum of 183 men and about 80 apples. The second run lasted 64 ticks or 73.8 seconds and had a maximum of 184 men and about 70 apples. And the third run lasted for 48 ticks or 58.8 seconds and had a maximum of 164 men and about 55 apples. The number of firms remained constant during all these runs, because the firms don't have a mechanism of getting out of the market. In this scenario, both men and firms have a big probabilityof reproduce and to produce, to be more precise 60%. When allowed to produce, the firms produce 4 apples that, if are not encountered by a man, have a lifetime of 5 ticks.

In the case of this scenario, even if the men and firms had good chances to reproduce and to produce, due to the fact that men did not know where the apples or firms were in order to go towards them when they did not have anymore energy there were only a few men that ate apples and got jobs. This sent the world directly into a crisis that ment that the men died, when they were left out of energy left, that only a small number of men were able to reproduce and that many of the apples were disused, because they simply dissapeared from the face of the world before being eaten.

The same thing happened for the second scenario too. As it can be seen in Figure 3, the three runs gave different results of time and population numbers even if they had the simillar inputs. Thus, the first run lasted 63 ticks or 73.8 secconds and had a maximum of 308 men and about 287 apples. The second run lasted 49 ticks or 58.8 seconds and had a maximum of 308 men and about 295 apples. And the third run lasted for 49 ticks or 58.8 seconds and had a maximum of 317 men and about 270 apples. The number of firms remained as well constant, but what is interesting is that during the first and second run, in the moment of the forth or fifth tick, the number of apples was bigger than the one of the men. In this scenario both the men and firms had good chances to reproduce and to produce (50%), the number of products was a high one (10 apples) with a rather small lifetime (3 ticks).

Compared to the first scenario, the second one had a bigger number of men that reproduced themselves (aproximative 50 men), but because of the same problems – men do not know where the firms and apples are in order to get more energy when they don't have anymore left – the world entered a crisis and the number of men soon decreased until there was no man left and the world stoped.

As it can be seen in Figure 4, a similar thing happened with the third scenario too. Only that here, another interesting thing happen: for all the three runs, the number of apples almost tripled compared to the number of men for exactly 3 ticks. Thus, the first run lasted 46.3 seconds od 45 ticks and had a maximum of 575 men an 1420 apples. The second run lasted 92.5 seconds or 80 ticks and had a maximum of 1440 apples and 504 men. And the third run lasted for 58.8 seconds or 55 ticks and had a maximum of 1270 apples and 500 men.

In the case of the third scenario, the number of men, firms and apples is the highest from all the three scenarios, but the chances a man has to repoduce are of only 30% and the one of a firm to reproduce are very high (80%) and the lifetime of an apple is quite small (only 2 ticks). This lead to a crisis of overproduction of apples and only to a small number of men. The men were able to reproduce themselves (approximatively 50 to 125 men), but soon died because they did not find a job or eaten apple in order to receive energy.

The simmilar results of the three scenarios demonstrate that our simple economy model can be improved by adding a way by which the firms can be paied with energy when an apple is eaten and another one by which firms can get out of the market.

The different results in population numbers of the three scenarios of our simple economy model, remind us that in fact everything is relative and that even if we give the same or similar inputs we can obtain very different results in time. The only thing we can do is adapt our solutions very well to the problem or crisis we encounter.

4. Acknowledgements

This article was written as part of the project "Doctorat în economie la standardele Europei Cunoaşterii (DoEsEc)" POSDRU/88/1.5./S/55287. Project co-financed by The European Social fund through POS-DRU 2007-2013 and implemented by the Academy of Economic Studies in partnership with the Western University of Timişoara.

5. References

- Ciutacu, Ileana, Dospinescu, Andrei (2011), The cost of exiting a system with connected benefits, article presented within The Sixth International Conference on Economic Cybernetic Analysis: Global Crisis Effects and the Patterns of Economic Recovery – GCER 2011, 20-21 May, Bucharest, Romania
- Farmer, J., Doyne, Foley, Duncan (2009) The economy needs agent-based modelling, Nature, 406, p. 685-686.
- Gaffeo, E., Gatti, D. D., Desiderio, S. & Gallegati, M. (2008), Adaptive Microfoundations for Emergent Macroeconomics, Eastern Economic Journal, 34(4), p. 441-463.
- Lengnick, Mattias (2011) Agent-based macroeconomics - a baseline model, Christian-Albrechts-Universität zu Kiel Economics Working Paper, 04, p. 1-28.
- Moisa, Cristina – Trends of the European Automotive Industry during the Economic Recession, Journal of Doctoral Research in Economics, Vol.II, Nr.1, Bucharest, Romania, 2010, pp. 3-11, ISSN: 2006-611X.
- Wilensky, U. (1997), NetLogo Wolf Sheep Predation model. http://ccl.northwestern.edu/netlogo/models/WolfSheepPredation. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanson, IL.
- Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

CHARACTERISTICS OF THE MOBILE COMPUTATIONAL ARCHITECTURE

CRISTESCU Marian Pompiliu
Associate professor Ph.D., Faculty of Economic Sciences, "Lucian Blaga" University of Sibiu, Sibiu, Romania, marian.cristescu@ulbsibiu.ro

CIOVICA Laurentiu
Junior teaching assistant, Ph.D. student, Economic Informatics Departments, Academy of Economic Studies, Bucharest, laurentiu.ciovica@gmail.com

MARTIN Florin
Junior teaching assistant Ph.D. student, Faculty of Economic Sciences, "Lucian Blaga" University of Sibiu, Sibiu, Romania, florin.martin@ulbsibiu.ro

Abstract: *Mobile devices have the necessary capabilities to provide services that assure security of the communication and integration of contextual information about the nodes (e.g. location, status, etc). Also, this service provides information about the QoS (quality of services) that allows devices to select the efficient communication method (e.g., SMS, GPRS, Bluetooth, etc). These devices are capable to access other devices and resources published by the Grid system, such as routing services that hold metadata and auxiliary information, or emergency services*.

Key words: mobile architectures, Grid systems, P2P nodes, topologies.

JEL classification: C61, C63, C88

1. Introduction

Mobile computing technologies as a generic term is to apply a wide range of portable, small, with wireless communication capabilities - as laptops with wireless technology, mobile phones, PDAs (Personal Digital Assistants) interfaces Bluetooth or IrDA (Infrared Data Association) - to provide access to information, communication and services at any location at any time, and by any means available.

Such mobile devices are transient network, connect and disconnects from the network frequently. Accordingly, the particular requirements of the applications based on mobility from the fact that these applications must typically be capable of operating in dynamic environments and unsafe / unstable and, in addition, considering cross heterogeneous operating systems, devices, host environments and services vary in different locations.

Technical solutions to achieve these objectives are often very difficult to implement because of mobile computing technologies require the creation of communication infrastructures that allow you to change network topology traditional operating systems and applications. In front of the lens to provide real mobility there are some constraints to be considered, which limits the ability to shift resources capacity compared to the fixed networks.

On the other hand, wireless devices (laptops, PDAs) which generally have limited resources (low processing power, battery life of the finite constraints of storage space), a major benefit when it becomes possible to use a considerable crowd of resources made available to other devices connected to network.

2. 'Static' and 'mobile' Grid nodes

Static Grid nodes provides the fallowing base capabilities, specific advantages of using Grid middleware GridLab (2007) to support the integrated architecture, and providing necessary support for the functionality of Mobile Grid Nodes in the network:

- tolerant/resistant to failure services: unlike the static Grid systems in which function the alignment and maintenance services of jobs waiting in line for execution, and in case of mobile Grid systems this is not possible because resources availability changes on a dynamic system (mobile devices, mobile Grid nodes, vary between connected/disconnected). Mobile nodes can disappear in the network unexpectedly due to disconnection of energy depletion, therefore it is necessary to continuously monitor service state MQS component, due to the fact that instable

mobile nodes do not diminishes substantial the performance. Strategy of establishing a maximum time interval for a job before shutting this to ensure resending a job when the result is not received in a predetermine time interval;

- hiding heterogeneity platform: one of the most complex problems to solve to create Mobile Grid architecture consists of hardware heterogenic resources and operating system (OS). To minimize costs to solve this problem we can use wide portability capabilities that allow rewriting Grid solutions in many languages and function on many different platforms;
- Decomposition jobs services: A job submitted for execution may contain a variety of overload due to the complexity of computing operations. Global load of a job can be too large to allow its processing by a single node in a static grid, within a reasonable time. In these conditions, the job must be broken down into several sub-jobs and send forward to execution to many participants in the static grid (e.g. in a complex operation or operations of simulation data warehousing). The problem becomes more difficult if the jobs initiated by mobile devices that do not have sufficient resources to carry out and complete those jobs (for example, searching for optimal routes between two points in the network in terms of multiple restrictions). Such jobs must be submitted for processing static grid system, which can complete the execution of jobs in a reasonable time;
- Allocation and planning jobs (scheduling): In some situations, when a job is sent for processing changes to static system Grid, can be distributed to a particular static node (or more static nodes) which are likely to hold most updated information necessary to execute the job, or computational capacity to process job in the shortest period of time. It is preferable that the jobs be assigned to the execution machine located closer to the location data that are necessary for the execution of jobs, thus can reduce network traffic and can reduce the scalability limits.
- Grid components with a planning role of jobs – called 'Schedulers' – react to the current availability of Grid resources. A 'Schedulers' automatically identifies the best machines that can run jobs that are waiting to be executed, and depending on the specific requirements of those jobs. 'Schedulers' components has the ability to calculate the approximate time of completion of the execution of a set of jobs are known in the case when machines used are fully "dedicated" to the static Grid system. If that would be used partly dedicated to the system resources and without an exact resources reservation, it exist the risk that resources unexpectedly become 'busy' – so cease to be available – and jobs can be suspended or delayed (delayed), such a situation can lead to complete every job unpredictable sites. If a component "Scheduler" has the role of 'broker' of resources and capabilities that involve 'bartering' associated capabilities 'scheduling'. An optimal scheduling system considers multiple resources, and this is a difficult math problem, such systems 'Scheduler' can use rules 'heuristics' (rules established to improve the likelihood of identifying the best combination of planning and booking of the jobs in order to optimize the outcome or any other indicator of the efficiency calculation);
- Resource reservation: Booking in advance of Grid resources for a given set of jobs. It was a step forward, after allocation and planning jobs ('scheduling') in the development of Grid system. Reservation may facilitate intervals set for completion of jobs and improve/guarantee the QoS (Quality of Services). When policies allow resources can be reserved in advance and partially dedicated to grid resources, or shared in the grid so that they run with low priority jobs for which they were high priority reserved.

These issues are addressed and the first implementations of Mobile Grid technology type. In addition, these technologies provide the capabilities necessary to identity mobile devices, security, interoperability and scalability of the system, using open protocols, standardized.

2.1. Architectural Option 2 - Mobile P2G, Mobile Grid integration and Peer-to-Peer

The second type of mobile computational architecture consists of a hybrid infrastructure called P2G (Peer-to-Grid), which integrates Mobile Grid P2P technologies. In this architecture, P2P techniques are used to organize mobile resources, grouping them in his peer group, authentication methods specific to this structure, management system group membership, group identity, etc.

The second type of architecture, described in this section is focused on the possibility of using 'light' technologies like J2ME Foster (2002) and KSOAP toolkit Foster (2002). These technologies can increase network capacity by exposure to services-based interfaces – used dynamic mobile nodes in the network – providing direct communication intra-Mobile-Grid – or between mobile nodes with a client connectivity using dynamic using peer-to-peer.

To accomplish P2G Mobile architectural model, activities are directed to:

a) – advancing mobile Grid technologies and integration of static and mobile Grid levels;
b) – advancing technologies Peer-to-Peer (eg JXTA Deelman (2005))
c) – integration of Grid technologies (and mobile Grid) with Peer-to-Peer technologies, allowing mobile nodes to communicate directly with each other and secured interface compatible with standard OGSA (Open Grid Services Architecture).

To facilitate this interaction, we aimed to define and use peer groups and 'light' instead of the conventional concept of Virtual Organization (VO) used in Grid systems in order to organize mobile nodes and infrastructure and provide the framework for implementing security policies for mobile nodes as defining the transitional concept 'and peer group'. This model progresses the P2P approach, providing connectivity intra – Mobile and Grid or between mobile devices involved in architecture.

3. Components of the Peer-to-Peer architecture

The proposed P2G architecture is an architecture consisting of three levels in which are connected the static grid level with the mobile network via peer-level Gateway components (the 'bridge' level). The three levels of architecture are:

- Static Grid locations/entities: static grid nodes;
- A group of mobile devices: peer's and peer's group;
- Gateway level consists of bridge, which interconnects the static and mobile nodes (resources)

Figure 1: Peer-to-Grid Architecture

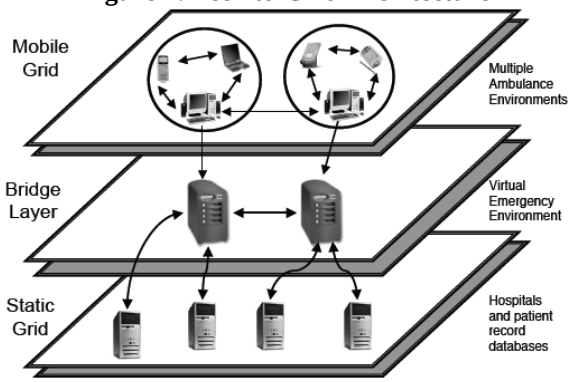


Figure 1 describes the three-tier architecture called Peer-to-Grid (P2G). The bridge level must have the capabilities necessary to meet the needs of communication in both directions, security requirements, services and transport infrastructure for the both static and mobile Grid levels and also for the peer's mobile network.

The second basic requirement is the need for mobile devices can share, aggregate and transmit data in a dynamic and unstable environment. It is notice that a decentralized approach to the mobile infrastructure is a more appropriate solution to meet this requirement, an architecture that allows mobile nodes to act as producers and consumers of services, events and messages – that is a Peer-to-Peer (P2P). The bridge level must integrate P2P network.

4. System and nodes typology

Diversity and transient nature of nodes requires a more flexible and dynamic mechanism of communication and discovery than in standard scenario and Grid infrastructure. Architecture mentioned requires P2P interactions. However, mechanisms need to communicate with other upwards Grid entities and a stable flow of information. Below is described how large static Grid configuration can interact with transitional networks and mobile P2P framework based in a hierarchical structure.

From a technological point of view, nodes from the system play four major roles:

1. Resources – represents the nodes (possible mobile nodes) which collect data. Those nodes have properties configured by other services and properties which can by obtain from them. Resources are characterized by "state", for example, have a geographic location, a property modification historic and, being connected or disconnected they are attached to a service or not.

2. Services - those nodes expose different functionalities of the network as a service. What these nodes offer can be actually a more 'passive' service, which means for example that they can consume

notification. WSRF services can expose resource-nodes as WS-Resources, offering a unified interface for a variety of resources. In the proposed model, resources can migrate between the services-nodes. If a resource is lost, then becomes necessary to be rediscovered. Service - nodes can communicate between through bridge – nodes.

3. Bridge Nodes – these nodes represents specialized services – nodes. These are capable to form a bridge level between the mobile level (service – nodes P2P) and static level (static Grid nodes). From the P2P network perspective, using bridge nodes as intermediary entities, static grid level can show entirely as another P2P node.

4. Grid Nodes – these nodes (seen as many nodes or, simple, as a single static grid nodes) represents a traditional Grid system with static IP address and strict infrastructure for security.

Communication between static grid level and P2p nodes it's made in both directions. Static grid is helped by P2P mobile collector nodes in collecting, extracting data, and static Grid analyze received data and return results.

Nodes with the role of receiving data represent a heterogeneous set of nodes in which some of them can play the role of communication hubs. As information is spread to the end of network, receiving nodes becomes more and more transient, more mobile and lightweight (for example mobile phones and laptops).

P2P topology allows the collecting and receiving nodes to communicate collectively, though the two types of nodes are very different from the point of view of the role they have in the network.

- without a P2P topology, in the first architectural version, all the nodes represents Grid server clients, and the server requires a connection to all the nodes; in such a typology the collecting nodes send data directly to the Grid system, and the Grid system further disseminate the data from the interested entities;
- using a P2P topology, the second architectural version, the collecting and receiving nodes can sustain themselves in maintaining the network robustness, spreading messages and resolving/translating the destination address.

Bridge level/node has as a primary function to transmit messages to and from static infrastructure. Can exist a big number of static Grid nodes which cooperate in the infrastructure, which can be organized in different Virtual Organizations with different intern politics 'intra-VO' and external cooperation politics 'inter-VO'.

5. Bridge node architecture

For the structural point of view, bridge level it presumes to be static and stabile, as the static grid level. Nodes which constitute the bridge level create an 'image' of the mobile network which the static Grid can understand and an 'image' of the static grid which P2P mobile network can understand. Internally, bridge nodes have the role to map and to translate between these two 'images'. The best solution in this case is considerer to be the proxy model.

Mobile nodes and peers groups are presented to Grid level through this proxy, allowing the Grid to understand and interpret as normal Grid entities. Base functionalities which the proxy must support are the fallowing:

- agent autonomy: Proxy can receive the role to take decision or to send messages in behalf of the name and in the benefit of a mobile node if this is not available and is necessary an immediate action. This function can require the cache capacity to hold information about state on the node or groups of nodes.
- message waiting line: to facilitate connectivity transparency, proxy must be capable to store messages received from the static Grid and to send them to mobile nodes when those become available again. This capability permits also the optimization of message transfer, allowing messages, for example the notifications, to be aggregated into one single message.
- translation or translation between contexts: an example of this kind of functionality appears between the security contexts. P2PS allows the nods and groups to be authenticated using P2PS certificates. Proxy can be member of the P2PS group and receiving the group messages. If one of these messages is requested by the static Grid, then the same proxy certificate can be used for authentication inside GSI (Grid Security Infrastructure) security system for message exchange with the static Grid.

6. Mobile node architecture

A mobile node is made from multiple modules which communicate with each other to facilitate the requested functionalities by the specific scenario.

1. Services level defines the interface with mobile node, respectively the protocols and the messages which the node can understand. This level is implemented as a WS-RF services interface. Exposed services by a nod can be divided in three categories: two times of services are shared by all the nodes (these are responsible services with migration service capabilities to another node and establishing data channels); the third tip are the services specific to the node, which depends by the nature and functionality of each node in the network.

2. SOAP Module parses and generate SOAP messages

3. Data channels model accept and generate usual binary data flux as images and audio/video streams. Details about connecting mode of this model are defined at the level of WS-RF services. A typical data channel can be described as a channel which accept (or wait) a certain MIME type, information's if the data flux is stream and where the data must be send forward (if this thing is necessary). The capability to establish / configure the 'in' and 'out' sources for a channel, allows the construction of data flux (workflows) in which data are send directly from node to node, without passing through a control node.

4. Migration module assures migration capabilities at the services level and any other resource from WS-Resource category hosted by the nod.

5. P2P module offers information's to services modules about the network vicinities and offers capabilities to discover another mobile nodes and translation the logical address. These capabilities are exposed by services levels as services compatible with the network.

6. Device model have the capabilities of introspect the device and return relevant information's which are requested by other modules. For example, can be necessary exposing the device location as an OGSA service or this location can be necessary to P2P module to decide the optimal transport modality to use for the connection to another node. Device model is also capable to describe the device as a WS-Resource. This resource is shown through an OGSA service interface.

7. Protocol model is responsible with different transfer protocol like HTTP or SIP (Session Initiation Protocol). Protocols can be selected directly by P2P level, or it can be sent by the service level or by the data channel module.

7. WSPeer

A generic implementation framework for P2P-Grid connectivity was created by the Cardiff University researchers. This generic framework, called WSPeer GEMSS (2003), has the potential to become relevant for any architecture which proposes them to integrate Mobile Grid and Peer-to-Peer technologies.

WSPeer is an API focused on double 'fertilization' of P2P technologies and Web/Grid technologies services. WSPeer use web services in an agnostic environment, making it possible to use standards like WSDL and SOAP in common scenarios, based on UDDI protocol for service discovery and on HTTP/THHPG protocol for transport, as in P2P oriented environment.

From the Grid service perspective, WSPeer incorporates also the capabilities of WSRF framework (Web Service Resources Framework) Deelman (2005), which allows use of services characterized by state or contextualization, as models of interaction of supported notification by WS-Notification. As a result, WSPeer can operate in many different conditions.

In addition to the capability to support simple web services interaction, WSPeer can interact with middleware Grid components using WS-RF framework and related specifications. In this way becomes possible an easy integration of different protocols in technological package of web services.

WSPeer support – in current form – a 'light' framework, independent from the domain, called P2PS, GEMSS (2003) – capable of publishing, advertising and communication in P2P ad-hoc networks. P2PS is a prototype initial implementation of the WSPeer generic framework. P2PS implements a subset of functionality of JXTA technology Foster (2002), using the abstraction of communication virtual channels ('pipe abstraction'), but customized for simplicity, efficient and stability. Through this, the WSRF-P2PS integration prototype allows services from the WSRF framework to be discovered and called in a P2P environment.

Similar to JXTA technology Foster (2002), P2PS uses the 'pipe' abstraction for defining communication channels. This virtual communication channels (pipes) can cross multiple transport ways which contain intermediary nodes. P2PS use this kind of logical addressing system and

translating/resolving services at terminal points for translating the logical address in specific network addresses. This process is useful especially for assuring the migration and mobility capabilities of hosts, allowing logical addresses to remain consistent while the location or the transport protocol modifies.

A mobile version of WSPeer is in the implementation process, called WSKPeer, representing a 'micro edition' of the WSPeer framework for mobile nodes. WSKPeer offer a subset of WSPeer functionality, WSDL parsing, WS-RF framework implementation support and Web services on mobile devices. WSKPeer is a mobile framework for SOAP messages exchange, created for devices with limited resources.

Using WSPeer allows implementing the base functionalities of the system nodes, in particular the capability to bind (like bridge model) different protocols and topology's and to function on devices with limited resource, simultaneously maintaining a service interface compatible with the standards.

Researching activities continues in direction of developing the WSPeer framework UDDI (2004) so will allow hosting services on mobile devices in Grid systems and generating events which can be consumed by services on other mobile nodes.

8. Conclusions

Global network typology is not necessarily found in the simple shape of a tree, but can be a complex typology with many entry points in client/server parallel networks. To maintain system scalability, any number of nodes can act as bridge nodes for static Grid level. These bridge nodes are referred using logical terminal points/address which are resolved/translate in the execution moment (runtime), exposing the static Grid as a peer node. In this way, multiple terminal bridge nodes can be used as Gateway components for static Grid level, in the case in which grid communities doesn't want to participate directly in the P2P network.

Offering a subset of WSPeer functionality, WSKPeer is in the developing process to allow the mobile devices to communicate between them and with the HTTP host using web services and message exchange in WS-RF. Combination between WSPeer and WSKPeer allows integration of static and mobile devices using an exchange system of standard and common messages.

9. References

- Altintas, I., Berkley, C., Jaeger, E., Jones, M., Ludascher, B., Mock, Kepler, S.: *An extensible system for design and execution of scientific workSSows*, In 16th Inter-national Conference on Scientific and Statistical Database Management (SSDBM), IEEE Computer Society, New York, 2004;
- Brookshier, D., Govoni, D., Krishnan, N., Soto, J. C.: JXTA: Java P2P Programming. Sams, March 2002;
- Deelman, E., Singh, G., Su, M.H., Blythe, J., Gil, Y., Kesselman, C., Mehta, G., Vahi, K., Berriman, G. B., Good, J., Laity, A., Jacob, J. C., Katz, D.: Pegasus: a Framework for Mapping Complex Scientific WorkSSows onto Distributed Systems, Scientific Programming Journal, (13 (3)):219–237, 2005;
- Deelman, E., Gannon, D., Shields, M., Taylor, I. J., Workflows for e-Science: An overview of workflow system features and capabilities, Journal of Future Generation Computer System, July 2008;
- Dynamic Data Driven Application Simulations, http://www.dddas.org/;
- Foster, I., Kesselman, C., Nick, J., Tuecke, S., The physiology of the grid: An open grid services architecture for distributed systems integration, Open Grid Service Infrastructure WG, Global Grid Forum, June, 2002;
- The GEMSS Project, http://www.ccrl-nece.de/gemss/index.html, 2003;
- GridLab: A Grid application toolkit and testbed, http://www.gridlab.org, 2007;
- UDDI Spec TC, UDDI Version 3.0.2, September 2004, http://uddi.org/pubs/uddiv3.htm.

EDUCATIONAL METHODS FOR INTELLIGENT TUTORING SYSTEMS IN ILL-DEFINED DOMAINS

CUCU Ciprian
Assistant professor, Ph.D., Faculty of Sciences / Department of Exact Sciences and Engineering, "1 Decembrie 1918" University, Alba Iulia, Romania, cucu.ciprian@gmail.com

CUCU Ioan
Associate professor, Faculty of Science / Department of Economical Sciences, University of Petroşani, Romania, cucuioan2002@yahoo.fr

Abstract: *In the development of e-learning, an increasing amount of attention is given to intelligent tutoring systems, i.e., systems that attempt to emulate the human methods and strategies of teaching. The current paper aims to show some of the ways in which educational methods can be implemented into such tutors, focusing on the most complex and challenging area – ill-defined domains, i.e., domains that lack a systematic solving approach.*

Key words: intelligent tutoring systems, ill-defined domains, education, e-learning

JEL classification: I23

1. Introduction

In 1984, Bloom proposed a series of studies that showed that one-on-one tutoring with a human expert tutor is the most efficient form of instruction. As these results remained unchallenged, researchers focused on solving a related issue: how to implement one-on-one tutoring for all learners in the context of limited resources of time, financial resources and human experience.

The development of Intelligent Tutoring Systems (ITS), i.e., systems that attempt to emulate the human methods of teaching, has been seen as the best possible answer to this challenge.

An ITS is usually defined as a system that has some type of knowledge representation of the domain it is teaching, is able to monitor the user's progress and predict the user's abilities and skills in that specific domain and adapts itself to the needs of the user (e.g. changing the sequencing of information provided, the speed or the level of details according to the users' abilities). (Murray, 1998)

The "traditional" architecture of an ITS is built on four components (Kodaganallur et al., 2005):

- the expert model: provides a representation of the domain knowledge, usually through production rules (but also other knowledge-representation methods);
- the student model: represents the student's knowledge at any given point, usually by comparing the student's answers to the expert model;
- the tutoring model: manages the relation between the student and the expert knowledge, deciding when and what remediation of the student's knowledge is necessary;
- the graphical interface: provides specific interaction tools.

Research into computer-based tutors that will behave similar to human tutors has continuously progressed and offered important results using two main paradigms: Model tracing tutoring and Constraint-based tutoring.

Model tracing tutors (MTT) are based on the ACT theory (Anderson et al., 1995), that aims to capture the process of reaching the solution to a problem. Such tutors use production rules to represent expert solution paths and then check the students' responses by comparing each student step with the steps from the expert solutions. MTTs have been implemented with good results (showing improvement in student results over traditional instruction) in a variety of domains including

a) College-level physics (Gertner & VanLehn, 2000; Shelby et al., 2001).
a) High school algebra (Heffernan & Koedinger, 2002).
a) Computer programming (Corbett & Anderson 1993).

Limitations of MTTs include the difficulty of development as well as some tracing errors: in some cases, correct yet unexpected student strategies (ie., that are not part of the expert model) can lead to erroneous remediation messages from the system.

Constraint-based tutors (CBT) are based on the works of Ohlsson (1994) - relying on a student's errors to build the student model. CBTs have also had various successful implementations in domains such as:

- SQL database commands (Mitrovic, 2003).
- Punctuation (Mayo, Mitrovic & McKenzie, 2000).
- Database modelling (Suraweera & Mitrovic, 2002)

CBTs are considered to be easier to develop, but also have more limitations related to misleading remediation and as well to the domains they can be applied to: they require domains in which the solution is rich in information, so that the remediation can be based on some strict connection between constraints and erroneous facts.

One important limitation of current ITS relates to the domains of applicability: while ITS have had good results in various fields (as shown above), all of these are so-called "well-defined" domains, relating to cognitive / problem-solving skills. Development of ITS in "ill-defined" domains, presented in literature (eg. Lynch et al., 2006), as domains that "lacks a systematic way to determine when a proposed solution is acceptable" is still in early stages and has yet to define specific architectures and models.

2. Ill-defined domains

As presented above, a tutoring domain (or knowledge domain) is considered ill-defined when there is no deterministic method proposing correct solutions in said domain.

There are various approaches to define the term 'ill-defined domain', but we consider the state of the art to be found in the characteristics proposed by Fournier-Viger et al (2010):

1. *Multiple and controversial solutions:* domains that do not have a clear solution determined by a predefined procedure, are ill-defined. This is, for example, the case with the domain law, in which arguments have logical connections, but no unambitious standard exists.
2. *No complete formal domain theory:* domains that do not have a structured, formal theory that can be applied to problems are ill-defined. Music composition is a perfect example, because even though there are a few theories related to the domain, they are incomplete and the subjectivity of evaluating aesthetics cannot be eliminated.
3. *Non-systematic task structure:* Analytical domains, in which tasks usually imply using incomplete and potentially incorrect information regarding a changing environment in order to make decisions , as well as „design domains" (Lynch et al., 2006), in which tasks involve creating new knowledge (eg. Composing music) are ill-defined.
4. *Abstract concepts:* Domains that work with concepts that are partially undetermined or do not have absolute definitions, are ill-defined, such as communication, where words or phrases can have different meanings, depending on the context.
5. *Overlapping sub-problems:* Well-defined domains have been shown to be susceptible to decomposition, i.e., complex problems can be divided in independent sub-problems that are easier to solve. Domains that can not be easily decomposed are considered ill-defined.

As a extension of the ill-defined domain theory, Fournier-Viger et al (2010) develop the theories of Aleven et al. (2008) and Lynch et al. (2006) and propose a set of characteristics for ill-defined tasks:

1. *The indefinite starting point:* The instructions or information necessary for solving the problem are incomplete or vague.
2. *The indefinite ending point:* The criterion that determines if the goal has been attained is complex and imprecise. There could be multiple and controversial solutions (
3. *Unclear strategies for finding solutions:* There are no clear strategies for finding solutions at each step of the problem-solving process.

3. Educational methods

- **Overview**

Building an ITS for an ill-defined domain is no trivial task, with challenges arising both in representing the knowledge of the domain as well as in defining the educational objectives and the strategies used by the tutor to 'guide' the learner in the right direction.

Research overview showed that there are three important methods that can be mixed in order to get the best educational outome: *interactive storytelling*, *mapping training objectives to content* and *user coaching*. (Lane et al., 2007)

- *interactive storytelling*

Interactive storytelling proposes that developing skills in ill-defined domains can be best achieved by secific psychology methods such as role- saying, that allows individuals to "experiment the desired behavior in a controlled environment where they can be onitored and can receive reinforcing or corrective feedback". (Segrin & Givertz, 2003).

Thus skills such as leadership, argumentation or abilities to create new knowledge (e.g. music composition) are best analyzed in various simulations: a music simulator might help analyze rhythm, while in a business simulation, leadership skills may be put to the test through communication with computer simulated characters.

The concept is based on the fact that experience is the best way to acquire tacit knowledge, fact that is the basis of pedagogical theories such as experiential learning (Kolb, 1984) and led to the development of several types of simulations, including game-based environments.

- *mapping educational objectives to content*

While the goal of the simulation, as stated before, is to allow the user a rich experience in the desired domain, some form of implementing eucational objectives is required in order to allow the system to analyise the behavior and draw conclusions based on the users' actions.

The simplest and most powerful way to do so, is to define specific 'milestones' in the possible interactions between the user and the system and use those as indicators of the presence of a certain skill (or at least part of a skill). For example, in the communication with a simulated character, if the user "rushes into business" i.e., does not make small talk to establish a report to the simulated character, this is an idication of a lack of skill.

Moving further, educational objectives should themselves be linked to eachother in ways that establish needed relations such as hirearchy (if objective B is met then there must be an objective A met before, otherwise there is a lack of skill from a user) or causality (objective A implies the necesity of objective B) to allow for more powerfull means of analysis.

- *user coaching*

In order to allow the user to develop an understanding of the domain and to conceptualize correct and incorect behavior in different contexts, the system must have a way of communicating with the user beyond the regular interface and interaction.

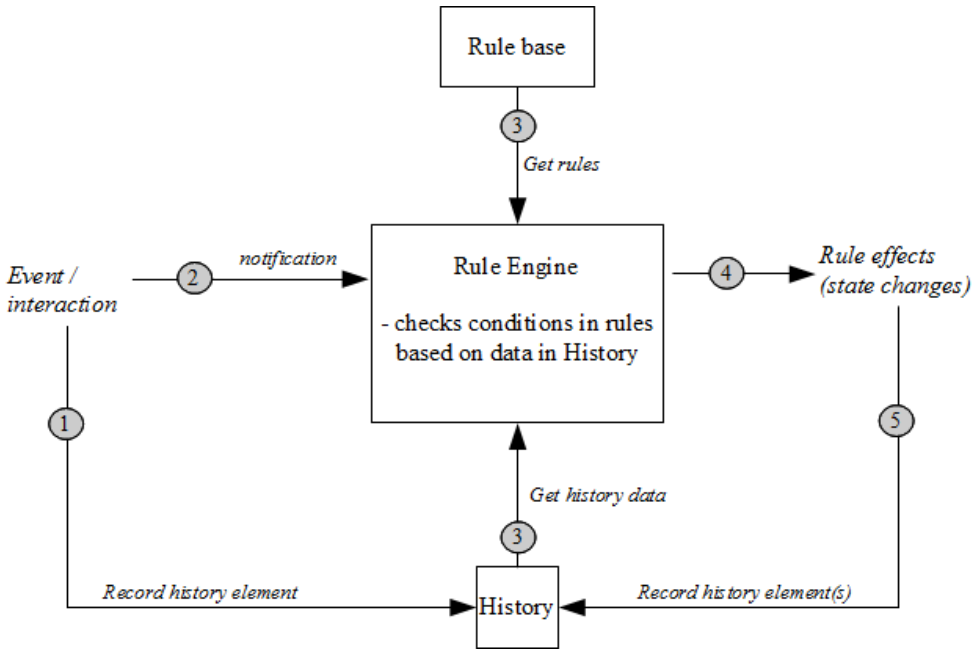
The user coaching can be implemented through a module that provides feedback on the users' actions, feedback that is linked to the educational objectives (as presented above). The coach "listens" to the user-system interactions and verifies the objectives according to the actions performed by the users. If the predefined relations between the objectives are satisfied at any given action, the user may be informed through a "reinforcement feedback"; if some relation is not satisfied, the user is given a "corrective feedback". The coach may also give the users hints i.e., forward-looking suggestions of what would be an appropriate next action, or warnings i.e., forward-looking suggestions of what should be avoided.

- **Case study**

To better illustrate the conception and implementation of educational objectives in an ITS used in an ill-defined domain, we present a case study on the educational component of the 'PM Game', a system developed in collaboration with the University of Zurich, that simulates social interactions from the perspective of project management scenarios, in order to analyse and develop social skills such as leadership and communication.

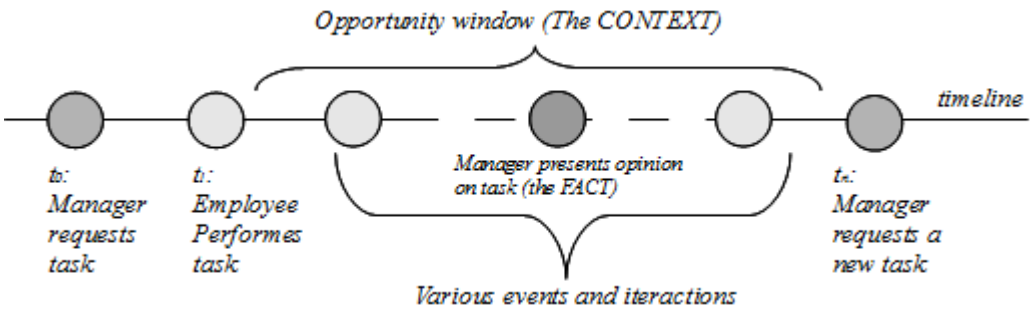
The system uses a custom rule-based expert model to implement domain knowledge. The customization consists mainly in a history module, that records all user-system interactions in order to use them to define the context of the interactions and to allow the system to understand the user's purpose in that given context. An overview of the rule engine and its connection to the history is available in Figure 1.

Figure 9: Recording and using the History in the PM Game



Furthermore, the actions performed by the users are evaluated in a specific time-frame (implemented for each action), that was named 'opportunity window'. This proposed that actions have a desired meaning only if they happen in this specific time-frame, otherwise they may have no meaning or even an opposing meaning. For example, a manager should provide his employees with feedback on the tasks performed, but should do so in a timely fashion, shortly after the task has been performed and before assigning a new task to the employee. Figure 2 shows how the PM Game implements the opportunity window for this use-case.

Figure 10: Opportunity window for feedback on employee task



The proposed educational method in the proposed system is customized to work together with the presented expert-model, offering positive feedback to the user when a positive behavior is observed, and remedial messages at negative behaviors.

As shown above, the production rules use facts from the history to define a context (window of opportunity) and to trigger some actions if a certain fact is found in the context. This applies to one instance, be it an instance of providing feedback to an employee, or one instance of asking a question to a customer (or other instances of social skills manifestation).

Taking into account more complex social scenarios, there is a need to define connections between such instances and thus create a *solution path* by checking if the instance that was captured by the rules happens in an appropriate moment (ie., if there are not other issues that should be addressed before it). For example, in a customer meeting scenario, one might consider the following path of interaction: 1. present yourself to the client → 2. ask question about his company → 3. ask questions about his needs →

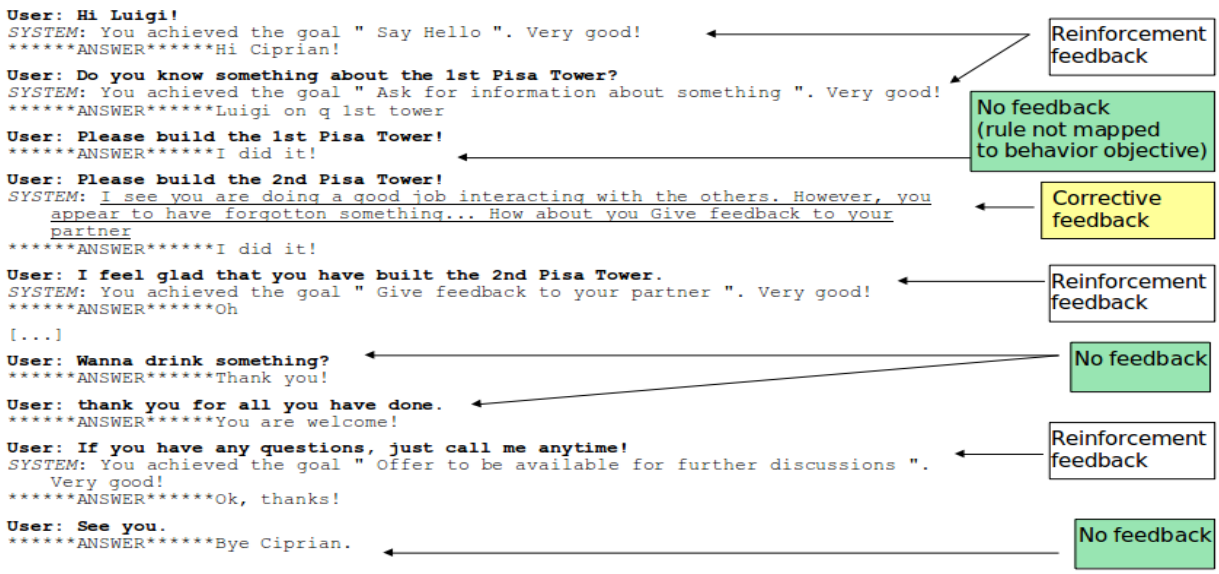
4. make a summary → 5. ask for a new meeting to present proposal. For each of these points, good behavior / bad behavior is defined using the expert model, but the order of the points should be on a different layer. On this layer, *goals* can be defined that can subsequently be mapped in the needed order for each scenario.

The design of the goal layer takes into account the following points:

- Goals are defined general and not specific to a certain scenario and a certain context. A good example is the goal "Ask about needs / expectations" versus "Ask about number of pages for website". The scenario designer can choose which goals and in which order they have to be applied for each scenario, thus creating the solution path for the scenario to suit his needs.
- Goals may be composed of one – many steps, that might also have an order if necessary. For instance a goal like "Present your company to the client" could have two steps - "Ask the client what he knows about your company" and "Give further information on your company to the client". The mapping between rules and goals can occur on step level, by a many-to-many relationship: the scenario designer should be able to choose any number of rules to map to any steps of the goals he selected in the scenario.
- The system allows both positive and corrective feedback to users, mapping each type of user-error to a specific error message indicating remedial possibilities. User-errors that are captured include:
  - Order error: the user fulfills a goal but leaves other previous goals unfulfilled; eg.: the user asks questions about the client's company but forgets to present his own company first.
  - Missed opportunity error: considering the concept of opportunity window defined for the expert model, when failing to act on an open opportunity (the window closes before fulfilling the rule), the user is informed with a specific remedial message. Eg.: forgetting to give feedback on an employee task, before requesting a new task.
  - Content error: this is a special category that allows to capture a behavior that's not at the correct moment, yet bypasses the order error (by not targeting any goal). For instance, consider the goal "Say hello" - that aims to check if the user always starts a conversation by greeting his partner. This would be a level 1 goal (ie., a goal that needs to be fulfilled first); if a user performs an interaction that triggers a different goal, the order error message (described above) would be shown to suggest remedial of behavior; however, if the user interaction triggers no goal, then the erroneous behavior (of not saying "hello" first) would not be captured. In this case, the goal itself ("Say hello") can be mapped to a set of rules that do not trigger the fulfillment of the goal, but trigger the "content error" message instead.

Figure 3 shows a part of a user-system interaction for a feedback use-case, emphasizing the feedback generated by the system.

Figure 11: Coach module example - 'feedback' use-case



4. Conclusions

The proposed paper gives an overview of the complexity of building intelligent tutors for ill-defined domains and focuses on the methods used to implement and evaluate educational objectives.

To this end, we propose three main elements - *interactive storytelling*, *mapping training objectives to content* and *user coaching* – that can be mixed to create a system that will constantly provide the user with feedback messages, either reinforcing his behavior (if it's considered good) or suggesting remediation (if it's considered bad). In this sense we provide proof both from research literature as well from our own work on the PM Game – a system that attempts to analyze and develop social skills such as leadership and communication.

For the near future, we wish to continue development of the tutoring module, in order to implement more complex relations between educational objectives (there is currently only a hierarchy relationship in place) and to complete the coach module (that currently provides only backward corrective or reinforcement feedback) with the ability to provide forward-looking messages, i.e., hints and warnings as defined in literature.

5. References

- Aleven, V., Ashley, K., Lynch, C., Pinkwart, N.: Proc. ITS for Ill-Defined Domains Workshop. ITS 2008, Montreal, Canada (2008)
- Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995), Cognitive Tutors: Lessons Learned. The Journal of the Learning Sciences , 4(2), 167-207.
- Bloom, B.S. (1984), The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13, 3-16.
- Corbett, A.T., & Anderson, J.R. (1993). Student Modeling In An Intelligent Programming Tutor. In E.Lemut, B. du Boulay & G. Dettori (Eds.) *Cognitive Models and Intelligent Environments for Learning Programming*. Berlin Heidelberg New York: Springer-Verlag.
- Fournier-Viger, P., Nkambou, R. & Mephu Nguifo, E. (2010). Building Intelligent Tutoring Systems for Ill-Defined Domains. In *Nkambou, R., Mizoguchi, R. & Bourdeau, J. (Eds.). Advances in Intelligent Tutoring Systems*, Springer, p.81-101.
- Gertner, A., & VanLehn, K. (2000). Andes: A Coached Problem Solving Environment for Physics. In G. Gauthier, C. Frasson & K. VanLehn (Eds.) *Intelligent Tutoring Systems: 5th International Conference, ITS 2000* (pp. 131-142). Berlin Heidelberg New York: Springer-Verlag.
- Heffernan, N.T., & Koedinger, K.R.(2002). An Intelligent Tutoring System Incorporating a Model of an Experienced Human Tutor. In S.A. Cerri, G. Gouardères, and F. Paraguaçu (Eds.): *ITS 2002, LNCS 2363*, pp. 596–608, 2002. Springer-Verlag Berlin Heidelberg 2002.
- Kodaganallur, V.; Weitz, R. R. & Rosenthal, D. (2005) A Comparison of Model-Tracing and Constraint-Based Intelligent Tutoring Paradigms. In *International Journal of Artificial Intelligence in Education* Volume 15, Number 2/2005.
- Lane H. C., Core M. G., Gomboc, D., Karnavat, A., & Rosenberg, M. (2007). Intelligent Tutoring for Interpersonal and Intercultural Skills. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC 2007)*, Orlando, FL.
- Kolb, D. A. (1984) Experiential Learning, Englewood Cliffs, NJ.: Prentice Hall
- Lynch, C. F.; Ashley, K. D.; Aleven, V.; Pinkwart, N.; (2006) - Defining "Ill-Defined Domains"; A literature survey.
- Mayo, M., Mitrovic, A., & McKenzie, J. (2000). CAPIT: An Intelligent Tutoring System for Capitalization and Punctuation. *International Workshop on Advanced Learning Technologies, IWALT 2000* (pp. 151-154). Palmerston North, New Zealand.
- Mitrovic, A. (2003). An Intelligent SQL Tutor on the Web. *International Journal of Artificial Intelligence in Education* 13, pp. 171-195.
- Murray, T. (1998), Authoring Knowledge-Based Tutors: Tools for Content, Instructional Strategy, Student Model, and Interface Design, in *The Journal of the Learning Sciences*, 7(1),5 - 64, Lawrence Erlbaum Associates, Inc.
- Ohlsson, S. (1994). Constraint-Based Student Modeling. In J. E. Greer, & G. McCalla (Eds.) *Student Modeling: The Key to Individualized Knowledge-Based Instruction* (pp. 167-189).

- Suraweera, P., & Mitrovic, A. (2002). Kermit: A Constraint-Based Tutor for Database Modeling. In S.A. Cerri, G. Gouardères & F. Paraguaçu (Eds.) *Proceedings of the Sixth International Conference, ITS 2002*, 377-387, Berlin Heidelberg New York: Spring-Verlag.

THE PHENOMENON OF INCREASING VALUE OF STOCK CAPITAL ON ROMANIAN MARKET

DOLCA Ioan
Faculty of Economic Science, University of Oradea, Oradea, Romania, dolcaus@yahoo.com

NICOLOV Mirela
Faculty of Economic Science, University of Oradea, Oradea, Romania, Faculty of Engineering, Aurel Vlaicu University, Arad, Romania, nicolovmirela@gmail.com

Abstract: *The purpose of this paper is to observe the influence of social capital on the Romanian stock market value. In this study was built a sample comprising 30 titles listed on the Bucharest Stock Exchange. The results obtained in this analysis is that social capital has a major influence on the value of the securities exchange. This can't be, because of the informational efficiency of the market, because everything in this analysis, this hypothesis was rejected. A possible explanation is the lack of information asymmetry, but due to the existence of large investors involved in the management or directors sometimes facing firms. In the case of the second problems, that the effectiveness of market information, results, based on the 30 titles studied, can not be exploited across the Romanian markets. Limitations of the study are data while the sample size and frequency of calculation of profitableness.*

Key words: Statistical Simulation Methods, Financial Econometrics , Financial Market, Analysis of Fiscal and Monetary Policy

JEL classification: C 15, C 58, D 53, E 63

1. Introduction

Economic researchers are concerned with different aspects of issuing shares. The role of information asymmetry in determining the issue price, investor reaction to announcements of the introduction or increase of capital stock, the reaction rates in an operation on a title belonging to a set of questions have to be determined.

Several studies have revealed a long-term underperformance of shares following a share issue, but not explain this behaviour or give them a sense of safety

2. Literature review

After the empirical study of Fama, Fisher, Jensen and Roll (1969), the impact on the value of a business operations involved in an event is measured by a study of event.

Assuming market efficiency in the semi-strong form, investors are able to anticipate at the creation or destruction value associated with an event, since the sale announcement. Therefore for measuring the variation of values of companies involved in an event, it is sufficient to calculate abnormal returns over a short title for a few days before the event and several days after the announcement event.

Some simulations carried out shows that for short-term, as few days, the sensitivity of results to the methodology used is very weak (Brown and Warner, 1980; 1985). Was shown that for long-term, over a period of several months, the results are very sensitive to the methodology used in the study of event (Fama, 1998).

The first stage of the event study methodology is to calculate abnormal returns of securities companies concerned. Abnormal return of a security is defined as the difference between observed profitability, profitability that was recorded, and return the theoretical one, as the return that would have been without event.

Was stated that short-term events studies, theoretical return securities index are estimated to use market model and the Capital Asset Pricing Model (CAPM), (Higson and Elliot,1998;Todea,2009; 2011). These methods require that profitability's titles are explained by a single factor: the market. At the same time, many empirical studies show that returns can explain the dynamics of securities based on other factors: size (Banz,1981), price-to earnings (P /E) (Basu, 1977; 1983) or book-to-market (Stattman,

1980; Lam 2002). Was pointed out that three factors are not independent, because the size, price-to-earnings and book-to-market are based on the same variable: price titles. In fact, the authors show that in the U.S. market, it is possible to capture the essence of dynamic yield securities during 1963-1990, based on two factors: size and book-to-market. In this context, was noted that are several approaches to calculate the enormous returns on long time (Fama and French, 1992). One is French model with three factors. Another method is to adjust the return of a portfolio return basis by reference. Reference portfolio consists of companies with similar business characteristics studied. And a third approach, enables the abnormal returns of securities that pass through an event title is the return adjusted by control of an enterprise profitability. Control firm is a business enterprise with similar features i.

Recent studies have used the temporal persistence of linear and, especially, nonlinear dependencies on the emergent stock markets (Hinich-Patterson, 1995). Were done studies for the Asian stock markets (Lim and Liew 2004) and for the Latin American stock markets (Bonilla et al., 2006). For the Central and Eastern European stock markets emphasized the existence of different stock price behaviors, namely long random walk sub-periods alternating with short ones, characterized by strong linear and/or nonlinear correlations. All these studies suggest that serial dependencies have an episodic nature, being also the main cause for the low performance of the forecasting models (Todea, 2011). An episodic behavior was identified on the currency markets also and it is attributed especially to certain political and social events(Brooks et al.,2000;2007; Lim and Liew , 2004;Hinich and Serletis, 2007; Todea, 2011)

3. Statistics and results

The purpose of this study is to observe the influence of social capital had the Romanian stock market value. For this was built a sample comprising 30 stocks listed on Stock Market Exchange Bucharest. The period studied for each title was one calendar year from the date resume quotation on the Stock Exchange securities from the own increase actual capital.

Abnormal return calculation was done by the difference between empirical returns, as that one which is registered, and returns the theoretical, predicted by the econometric model:

$$R_{it}=\alpha_i+\beta_i R_{Mt}+\varepsilon_t \quad (1)$$

Where Rit – title and theoretical returns for i-th ;

RMt- market returns, illustrated in our case by the BET-ROL.

For the specified model we used time series consist of returns on securities registered in the case in a period of one year before the disruption announcement and the stock market quotation, ie BET-ROL in the same period.

To estimate the parameters $\hat{\alpha}_i$ and $\hat{\beta}_i$ using the method of least squares .

Compliance testing was performed classic mismatch hypothesis residue of normality and their homoscedasticity. Results are displayed below.

Residues of the improper assumption was tested using the Durbin-Watson test for correlation of order 1 and with Ljung-Box test for higher order correlation.

Durbin-Watson test involves:

$$d = \frac{\sum_{i=2}^{n}(\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1})^2}{\sum_{i=1}^{n}\hat{\varepsilon}_i^2} \quad (2)$$

Q test of Ljung-Box statistic is a statistical test of the null hypothesis of lack of correlation up to order k with its corresponding p's and is designed as:

$$Q_{LB} = T(T-2)\sum_{j=1}^{k}\frac{r_j^2}{T-j} \quad (3)$$

where rj is the autocorrelation number j and the number of observations.

Asymptotic Q is distributed as a $\chi 2$ with a variable number of degrees of freedom equal to the number of autocorrelations.

Hypothesis of normality of residues was tested with Jarque-Bera test, which is a very sensitive test based on a variable $\chi 2$ with 2 degrees of freedom, of the form:

$$JB = \frac{N-k}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right) \quad (4)$$

Where k-number of estimated coefficients used to create the series,
K-asymmetry coefficient,
S-coefficient of vaulting.

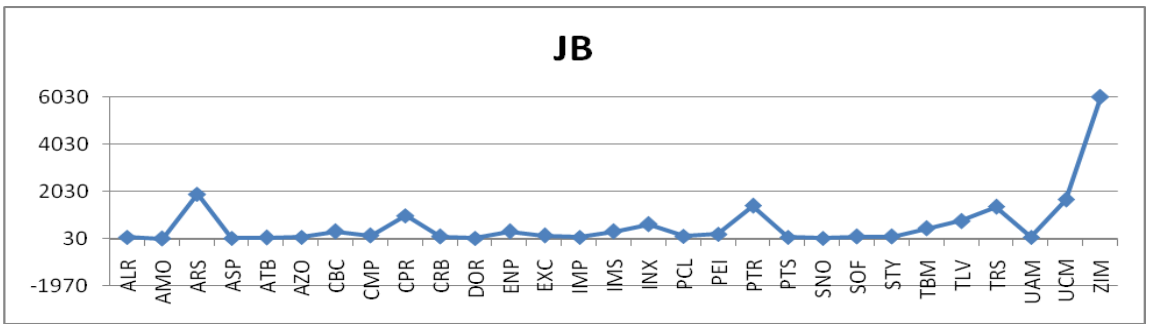
Results of the Durbin Watson test are listed in table 1 .

Table 1: List of companies used in the study and Durbin-Watson test results

	companies	D-W	test results	Q(25)	p	test results
1	ALR	1,424847	yes	72,164	0	yes
2	AMO	1,810015	no	21,26	0,678	No
3	AMP	1,825957	no	10,546	0,995	No
4	ARS	2,215131	no	30,328	0,212	yes
5	ASP	2,147363	no	34,722	0,093	yes
6	ATB	1,347374	yes	97,963	0	yes
7	AZO	2,248882	no	28,877	0,269	yes
8	CBC	1,799403	no	21,978	0,637	No
9	CMP	2,201001	no	42,151	0,017	yes
10	CPR	1,990123	no	24,906	0,468	yes
11	CRB	1,748391	no	27,522	0,33	yes
12	DOR	2,005818	no	25,203	0,451	yes
13	ENP	1,578143	yes	49,277	0,003	yes
14	EXC	2,082635	no	19,209	0,787	no
15	IMP	1,883520	no	24,525	0,489	yes
16	IMS	1,754416	no	34,963	0,089	yes
17	INX	1,976422	no	33,946	0,109	yes
18	PCL	1,851812	no	30,782	0,196	yes
19	PEI	1,579721	Yes	45,077	0,008	yes
20	PTR	1,549950	Yes	45,398	0,008	yes
21	PTS	2,101940	no	22,002	0,636	no
22	SNO	2,285060	no	30,472	0,207	yes
23	SOF	1,988353	no	32,684	0,139	yes
24	STY	2,211223	no	35,808	0,075	yes
25	TBM	1,756654	no	81,094	0	yes
26	TLV	1,337394	yes	57,612	0	yes
27	TRS	1,782987	no	21,3	0,676	No
28	UAM	2,234400	no	19,759	0,759	No
29	UCM	1,509569	yes	55,177	0	No
30	ZIM	1,900151	no	13,04	0,976	no

Source:Authors calculus

Figure 1: Results for Jarque-Bera Test . For all the used companies the value for p is zero.



Source:Authors calculus

We note, with the test that does not comply with any residue series and asymptotic normality assumption given the large number of observations. We conclude that residues tend to follow a normal law.

The hypothesis was tested with White Test for homoscedasticity. This is a test for heteroscedasticity of the normally residues of regression. Through parameters estimated by least squares method , values are valid in this heteroscedasticity meaning, but the standard errors obtained are not valid.

White test is a test of the null hypothesis of no heteroscedasticity state, the alternative to the heteroscedasticity existence .Test is based on auxiliary regression are introduced residues squares regressions all cross-products.

For example, suppose we estimated regression:

$Y_1=b_1+b_2x_2+b_3z_t+\varepsilon_t$ (5)

The test is based on auxiliary regression

$\varepsilon_t=\alpha_0+\alpha_1x_t+\alpha_2z_t+\alpha_3x_t^2+\alpha_4z^2_t+\alpha_5x_tz_t+v_t$ (6)

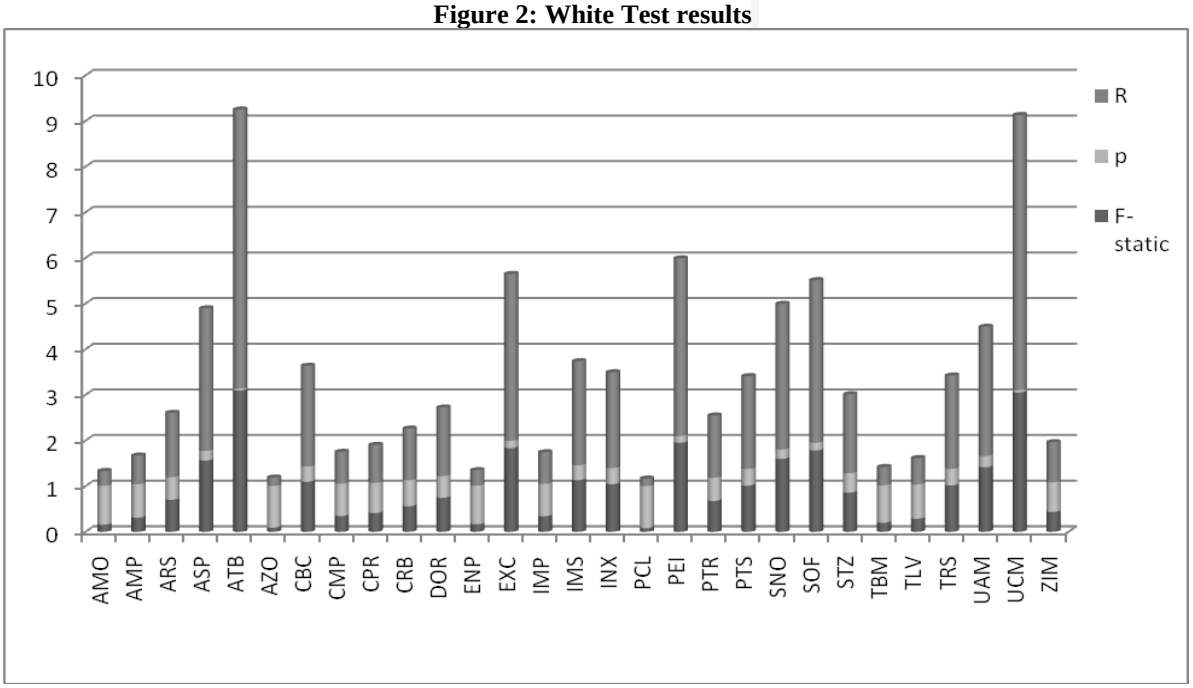
The results of White test for the studied companies are presented in Table2 and Figure 2

Table 2: White Test

	Societăţi	F-static	p	R	p
1	**ALR**	**30,17591**	**0**	**49,09017**	**0**
2	AMO	0,158614	0,853413	0,320668	0,851859
3	AMP	0,311048	0,732966	0,628069	0,730494
4	**ARS**	**0,700228**	**0,497455**	**1,409475**	**0,494238**
5	**ASP**	**1,562488**	**0,211678**	**3,123415**	**0,209778**
6	**ATB**	**3,09445**	**0,047061**	**6,110951**	**0,0471**
7	AZO	0,091745	0,912369	0,18558	0,811385
8	**CBC**	**1,098476**	**0,335002**	**2,204032**	**0,332201**
9	CMP	0,346237	0,707689	0,698952	0,705058
10	CPR	0,411439	0,663149	0,830108	0,660305
11	CRB	0,561108	0,571302	1,13071	0,568159
12	**DOR**	**0,74616**	**0,47525**	**1,501374**	**0,472042**
13	ENP	0,167293	0,846048	0,338191	0,844428
14	**EXC**	**1,832717**	**0,162147**	**3,655704**	**0,160759**
15	IMP	0,342237	0,710516	0,690874	0,707911
16	**IMS**	**1,136515**	**0,322608**	**2,279658**	**0,319874**
17	**INX**	**1,045058**	**0,353221**	**2,097751**	**0,350331**
18	PCL	0,080939	0,922274	0,163737	0,921393
19	**PEI**	**1,95372**	**0,14393**	**3,893308**	**0,142751**
20	PTR	0,676483	0,509341	1,36194	0,506126
21	**PTS**	**1,013498**	**0,364451**	**2,034916**	**0,361513**
22	**SNO**	**1,598053**	**0,204374**	**3,193601**	**0,202544**
23	**SOF**	**1,783829**	**0,170151**	**3,559576**	**0,168674**
24	**STZ**	**0,860838**	**0,424072**	**1,730525**	**0,420941**
25	TBM	0,199029	0,819659	0,402295	0,817792
26	TLV	0,287068	0,750711	0,579761	0,748353
27	**TRS**	**1,017498**	**0,363009**	**2,042881**	**0,360076**
28	**UAM**	**1,415102**	**0,244863**	**2,832127**	**0,242667**
29	**UCM**	**3,052830**	**0,049011**	**06,030742**	**0,049028**
30	ZIM	0,436584	0,646738	0,88066	0,643824

Source:Authors calculus

Figure 2: White Test results



Source:Authors calculus

Figure 2 shows the corresponding titles of the model whose residues are heteroscedastics. Autocorrelation and heteroscedascity residues, influence the quality variation and covariation matrix which estimate of model parameters. The correction method in this situation give the general case of an econometric model linear general form (written as matrix): $Y = AX + \varepsilon$,

The estimation matrix variation and covariation the estimators in the presence of autocorelation and unknown degree of heteroscedasticity, as the relationship:

$$\Sigma_{NW}=(XX)^{-1}\hat{\Omega}(XX)^{-1} \quad (7)$$

where:

$$\hat{\Omega} = \frac{T}{T-k}\{\sum_{t=1}^{r}\varepsilon_t^2\, x_t x_t' + \sum_{v=1}^{q}\left(1-\frac{v}{q+1}\right)\sum_{t=v+1}^{r}(x_t\varepsilon_t\varepsilon_{t-v}x_{t-v}' + x_{t-v}\varepsilon_{t-v}\varepsilon_t x_t')\} \quad (8)$$

k-the number of regressors, lag q is truncated and the number of autocorelation used in assessing dynamics εt residues. authors aimed at determining the relationship of q as q=4(T/100)2/9. Based on this correction, the data and values will be corrected by Student static relationship

$$t_{NV} = \frac{\hat{\beta}_j}{\sqrt{\hat{\mathbb{V}}_{NW}(\hat{\beta}_j)}} \quad (9)$$

where $\hat{\mathbb{V}}_{NW}(\hat{\beta}_j)$ is the corresponding element βj of matrix variation and covariation ΣNW. In the case of homoscedastics models with uncorrelated errors, we will only estimates only the values tcalc of the parameters accordingly.

Aggregation of abnormal returns

At this stage, abnormal returns were calculated for each individual and cumulative abnormal return on buy-and-hold, using the formulas:

$$RAC_{i[a,b]} = \sum_{t=a}^{b}[R_{it} - E(R_{it})] \quad (10)$$

$$RABH_{i[a,b]} = \prod_{t=a}^{b}[1+R_{it}]^t - \prod_{t=a}^{b}[1+E(R_{it})]^t \quad (11)$$

Table 3: Cumulative abnormal return for each titleand return on buy-and-hold abnormal

	companies	RAC	RABH
1	ALR	-0,18868	-6,1E+56
2	AMO	-0,49616	-2,4E+21
3	AMP	-1,9677	-1E+153
4	ARS	-1,16773	-5,4E+74
5	ASP	-0,83124	-7E+53
6	ATB	-1,28751	-5E+79
7	AZO	0,077109	-8E+42
8	CBC	-0,32368	-2,9E+43
9	CMP	-0,58123	-1,7E+55
10	CPR	-0,59568	-3E+16
11	CRB	-1,9518	-604E+86
12	DOR	-0,71792	-2,6E+73
13	ENP	3,57579	-2E+118
14	EXC	0,1511786	6,73E+64
15	IMP	0,623995	1,85E+13
16	IMS	-0,27529	-70612,5
17	INX	-0,54296	-2E+44
18	PCL	0,10201	6,18E+41
19	PEI	-0,25022	-1,2E+55
20	PTR	-1,61022	-5,3E+38
21	PTS	-1,746237	-2,6E+23
22	SNO	-0,46544	2,58E+32
23	SOF	-0,75939	-1,4E+53
24	STZ	0,189663	7,54E+57
25	TBM	1,431147	9,01E+22
26	TLV	-0,56598	-4,5E+46
27	TRS	-1,93581	-9,6E+27
28	UAM	-0,44446	-5,5E+50
29	UCM	-0,96617	-8,15452E+60
30	ZIM	0,137517	-2,8E+15

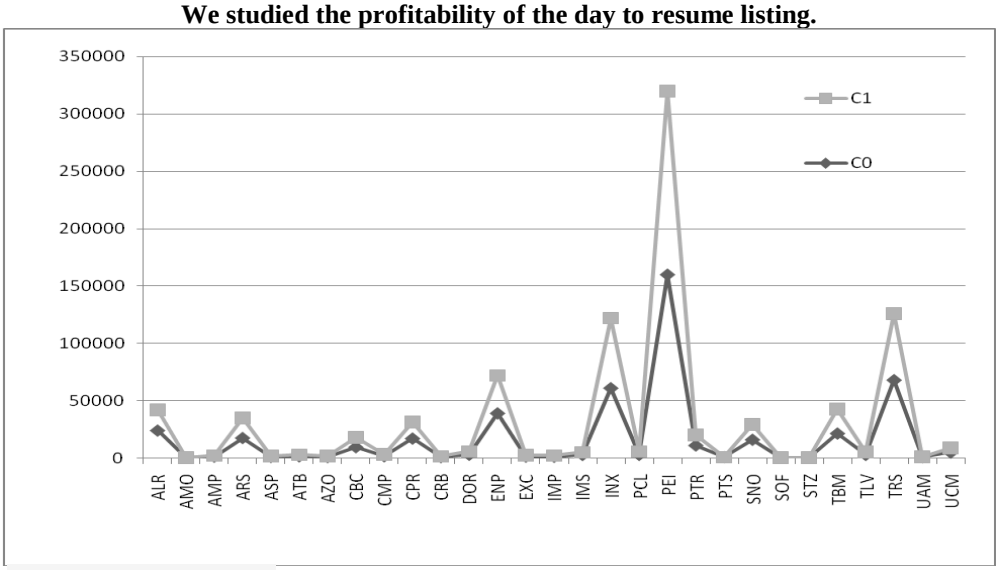
Calculating the average abnormal return we obtaine: $\overline{RAC}_{[a,b]}$ = -0,65227 and $\overline{RABH_{[a,b]}}$ = -3,8041E+151
Testing the importance of abnormal returns
To test the value of the calculated significantly:

$$t = \frac{\overline{RA_{[a,b]}}}{\sqrt{\sigma^2(RA_{a,b})/N}} \quad (12)$$

We will compare the value calculated tabulated the total amount of t for α = 0.05 and v = 30, ie t tab = 2.042.

Let's make the Hypothesis:
H0: the average abnormal returns of the N securities is zero,
the alternative
H1: the average abnormal returns of the N securities is other than 0.
If cumulative abnormal return is tcalc = -0.124, in this case, since | t calc | <t tab, we accept the null hypothesis which reject the contrary, that we can say with 95% accuracy that the average profitability of the securities is normal zero. Thus increasing the capital of companies made no demand, no damage value of the securities exchange.
If abnormal return on buy-and-hold is tcalc = -0.171, in this case, since | t calc |<ttab, we accept the null hypothesis and is rejected the contrary situation. We can say with precision, that the average returns of 95% the normal titles is zero. Increasing the capital of companies ,made no demand, no damage value of the securities exchange.
For the second objective of the study, namely that of determining the type of informational efficiency of the Romanian market, we studied the profitability of the day to resume listing.
For informational efficiency hypothesis to be accepted, we must give the average returns, which are calculated as zero, when later today , these reflect daily course from tomorrow . If this difference is not zero, we can say that the Romanian market is inefficient as information.

Figure 2: The type of informational efficiency of the Romanian market.
We studied the profitability of the day to resume listing.



Source: Authors calculus

$\bar{R}_i$= - 0,08761 to test the hypothesis that the average student Test is equal to 0. Tcalc obtain a = -3.474126
We observe that the probability of accepting the null hypothesis is 0.0017, thus rejecting the null hypothesis is accepted as contrary.
The average is significantly different from 0. Since Ri ≠ 0, informational market efficiency hypothesis is rejected and is accepted the null alternative hypothesis. So the Romanian market is informational efficient.

4. Conclusions

The results of such analysis is highly dependent from the chosen methodology. If, for example, instead of the market model would be taken as reference returns of a portfolio of securities of companies with similar characteristics studied business or for reference for field operating company, the results would have been totally different.

The results obtained in this analysis, as shown, social capital has a major influence on the value of the securities exchange. This can't be but on account of the informational efficiency of the market, because everything in this analysis, this hypothesis was rejected.

A possible explanation is the lack of information asymmetry, the existence of large investors involved in managing or dealing sometimes with companies administrators.

In the case of the second issue, that of market information efficiency, results, based on the 30 titles studied, can not be exploited across the Romanian market. These results fit line results in all major global markets. Limitations of the study are data while the sample size and frequency of calculation of returns. It would be interesting to do such a study for all companies listed on BSE, extending the findings to the entire market.

Methodology proves that the important excess returns are found especially in subperiods with linear and nonlinear correlations. Therefore, we can say that profit opportunities appear from time to time and that we are not dealing with a continuous improvement of the informational efficiency degree of the Romanian currency market, as predicted by the Efficient Market Hypothesis (Todea , 2011). Previous results plead in favor of the Adaptive Market Hypothesis developed by Lo in 2004 rather than the Efficient Market Hypothesis. Conform to the evolutionist principles of new approaches, profit opportunities in the Romanian currency market are manifesting themselves in an episodic way and they do not disappear definitively. A similar conclusion is revealed for the North-American market also by the study conducted by Neely et al. in 2009, which concentrated on a series of classical studies of the specialized literature (Todea , 2011).

5. References

- Banz R.W., (1981), The relationship between return and market value of common stocks, *Journal of Financial Economics*, vol.9, Issue 1, March 1981, p.3-18.
- Basu S., (1977), Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis , The Journal of Finance , vol. 32, No.3, p.663-682
- Bonilla, C.A. Romero-Meza, R. and Hinich, M.J. (2006), Episodic Nonlinearity in Latin American Stock Markets Indices, Applied Economics Letters, 13, p. 195-199.
- Bonilla, C.A. Romero-Meza, R. and Hinich, M.J. (2007), GARCH inadequacy for modeling exchange rates: Empirical evidence from Latin America, Applied Economics, 39(19), p. 2529-2533.
- Brown S., Warner J.B., (1985), Using daily stock returns: The case of event studies, Journal of Financial Economics , vol.14, Issue 1, March 1985, p.3-31.
- Cosh A. , Hughes A., Takeovers After Takeovers, Working Paper forms part of the CBR Research Programme on Corporate Governance, 2008, document accesed april 2012 : 10.1.1.167.5920.pdf
- Fama E. F., (1998), Market eficiency, long-term returns, and behavioral finance, Journal of Financial Economics vol.49, Issue 1, p. 283-306.
- Fama E., French K., (1992) , Cross Section of Expected Stock returns , The Journal of Finance, XLVII, nr2, p. 427-465.
- Fama E.F., Fisher L., Jensen M.C. and Roll R.W., „The Adjustment of Stock Prices to New Information" (February 15, 1969). Reprinted in International Economic Review, Vol. 10, February 1969; Strategic Issues In Finance, Keith Wand, Ed., Butterworth Heinemann, 1993; Investment Management: Some Readings, J. Lorie, R. Brealey, eds., Praeger Publishers, 1972. Available at SSRN: http://ssrn.com/abstract=321524 or doi:10.2139/ssrn.321524, http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1468-2354
- Florea Ioan (2003),Econometrie, Oradea University Printing House.
- Florea Ioan, Parpucea Ioan,Buiga Anuța, Lazăr Dumitru (2000) Statistică inferențială –Cluj University Press.
- Higson A., (2003), Corporate Financial Reporting, Theory and practice, Sage Publications Ltd., London.
- Higson, C. and J. Elliott (1998), Post-takeover Returns: The UK Evidence, Journal of Empirical Finance, 5, p. 27–46.

- Hinich M.J. and Serletis A. (2007), "Episodic Nonlinear Event Detection in the Canadian Exchange Rate", Journal of the American Statistical Association, 102, p. 68-74.
- Hinich, M.J. and Patterson, D.M. (1995), Detecting epochs of transient dependence in white noise, Mimeo, University of Texas at Austin Press.
- Lam K.S.L., (2002), The relationship between size, book-to-market equity ratio, earnings–price ratio, and return for the Hong Kong stock market, Global Finance Journal , vol.13, Issue 2., p.163-179.
- Lim, K.P. and Liew, V.K.S. (2004), Non-linearity in financial markets: evidence from ASEAN-5 exchange rates and stock markets, ICFAI Journal of Applied Finance, 10(5), p. 5-18.
- Loughran T.,Ritter J. (1995) The New Issues Puzzle, The Journal of Finance, vol., 50, Issue 1, (Mar.1995), p.23-51.
- Stephen J. Brown, Jerold B. Warner, (1985) Using Daily Stock Returns ,The Case of Event Studies, Journal of Financial Economics, 14, p. 3-31.
- Todea A., Silaghi S., (2001), Legea statistică a rentabilităţilor activelor cotate la Bursa de Valori Bucureşti; între legea normală şi alte legi stabile, Studia Universitatis Babeş-Bolyai, Oeconomica, Cluj-Napoca, XLVI, 1, p.129-136.
- Todea A., Tulai H., Plesoianu A.,(2009), D-CAPM: Empirical Results On The Bucharest Stock Exchange , Theoretical and Applied Economics, Vol.12 Issue 12, p.632-639.
- Todea A., Zoica-Ienciu A., (2011), Technical Analysis and Stochastic Properties of Exchange Rate Movements: Empirical Evidence From The Romanian Currency Market, Romanian Journal of Economic Forecasting , vol. 1/2011, p.175-192.
- Todea Alexandru , Caracterul instabil al riscului de piață evidențiat empiric pe piața românească, on site-ul www.kmarket.ro accesed in octomber 2011.
- Todea Alexandru-(2002) -Teoria pieţelor eficiente şi analiza tehnică:cazul pieţei româneşti, Studia Universitatis Babeş Bolyai,Oeconomica,XLVII,1.
- Tuch C., O'Sullivan N., (2007) , The impact of acquisitions on firm performance: A review of the evidence, International Journal of Management Reviews, vol.9, Issue 2, p.141-170.

IMPLEMENTING STRATEGY: THE KEY FACTORS WHICH CAN IMPROVE THE SUCCESS RATE OF INFORMATION SYSTEM DEVELOPMENT

GHEORGHE Gabriela
Ph. D., Dunarea de Jos University of Galati, Romania, gabriela.gheorghe@ugal.ro

LUPASC Adrian
Ph. D. Assistant professor, Dunarea de Jos University of Galati, Romania, adrian.lupasc@ugal.ro

Abstract: *The aim of this paper is to highlight the role of the deployment of the Informational System in reaching the organization's major goals, by presenting the risks and positive effects of the IS modernization, exemplifying by a public educational organization study case. The article will analyse the importance of the organizations' focusing on reconsidering methods and decisional techniques as integrated elements in the informational system having as main criterion enabling competitive edge of the organization.*

Key words: informational system, decision-making process, business technology applications, ERP, PERT

JEL classification: O 33, L 21, D 78

1. Introduction

In the current economic environment, characterized by competition and rapid evolution, the informational system itself must adapt to this evolution in order to offer real time information and facilitate decision-making. The adaptation of the decision-making process can make the difference between a prosperous organization and one in deficit, between a public service or an efficient administration and a bureaucratic one.

The reaction speed of the Informational Systems is essential in keeping the pace with the current economic changes. Sharing information among different actors is possible via new technologies like virtual community and web 2.0.

Enterprise resource planning (ERP) is an information system (IS) that integrates and manages all business aspects of the business knowledge, including production planning, purchasing, manufacturing, sales, distribution, accounting and customer services. (Jacobs, F.R; Bendoly, E, 2003)

Business Intelligence tools can be interfaced to the ERP system to improve the decision making of managers and provide useful knowledge relative to enterprise positioning, market trends and information on competitors. (Parker, B., 2011)

The use and adoption of ERP system depends on many factors both internal and external. This qualitative study aims at identifying the key factors for successful implementation of "Enterprise applications for running your business". (Hamerman, P.D, 2011)

2. The premises of Informational System implementation

The implementation of an informational system is an important stage for an organization. According to the 2009 study conducted by Forrester Research on a number of 2200 IT decisional factors in Europe and the U.S., the nr. 1 priority is the modernization of existing key informatics systems. (Ried, S., 2009)

Modernization is referred to as regrouping in a single system for all processes and operations for a unified leading of the activity, in order to avoid dysfunctionalities linked to the interface and updating in the case of existing separate systems for different activities. In the case of multinational firms, with branches or based on distribution networks, the stake of the unification of the workflow simultaneously with the integration of the Information System is the reason of the decision to modernize the SI.

The ERP market recovered fairly well in 2010 from the midrecession activity levels of 2009; the market will continue to grow in 2011, but will gradually shift from a licensed to a subscription model over the next five years. (Martens, C.; Hamerman, P.D., 2011)

In an age where all business technology applications should provide some competitive advantage, according to the Forrester's survey results, 76 percent of IT chiefs said that the operational efficiency was the goal of their organization's ERP investment.

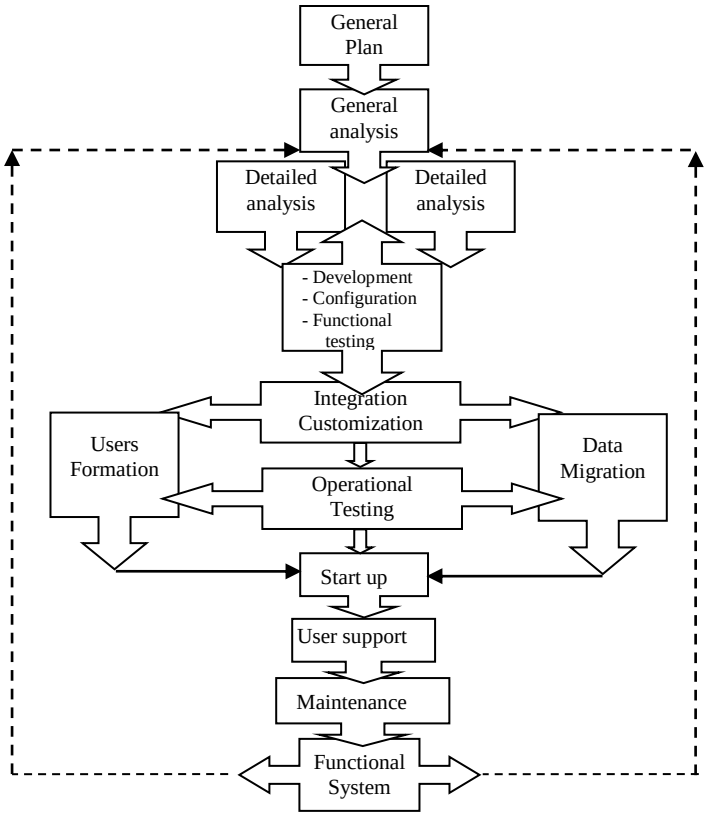
There are some industry differences as to when you can have competitive advantage with ERP. For example, within the financial services industry, "everyone's got it, so there's not really competitive advantage". But in the real estate industry, a robust ERP package that is embraced by users can be more of a differentiator because they aren't as common in this sector as they are in industries such as financial services and manufacturing. Then it becomes clear how efficient you are at using your ERP systems. (Wailgum, T., 2008)

"The use of SAP Business Intelligence shows how the creation of a culture of information democracy contributes directly to increasing the performance. Companies that use Business Objects solutions have long overcome the phase of a simple access of or information providing, offering their own employees, at all the levels of the organization, to make the most appropriate decisions as fast as possible, which means increased profitability for the organization", declared Thomas Jeandet, Regional Presales Manager for Eastern Europe, SAP Business Objects.

3. Masterplan of an ERP type IIS implementation

The implementation sequence in a multistage approach is presented in the next figure:

Figure 1: Stages of an IIS implementation



Source: Adapted from SAP implementation methodology

Stages of the life cycle of a project for implementing an Integrated Informational System (Niculescu, O.; Plumb, I.; Pricop, M.; Vasilescu, I.; Verboncu, I., 2004)

1. Identifying the need for implementing an information system, analysis of possible solutions, optimal decision making

Phases of the decisional process:

1.1 Defining the decision situation on implementation of an IS

1.2 Detailed defining of the project

- technical specifications
- criteria for assessing the success of the project
- risk analysis, assessment of probability of occurrence of events that may affect the project, providing for actions to reduce, eliminate or sharing risks right from the early stages of the project

1.3 Selection grid: Identification of the most important criteria for the process and the final result and percentage setting for each criteria weight
1.4 Analysis and evaluation of all possible solutions
1.5 Calculating the consequences of these solutions and of the possibility of satisfying the established criteria
1.6 Choosing the best option
2. Strategy elaboration, goal setting, general planning and change management, communication
2.1 specialists' team setting
2.2 project team informing and training
2.3 informing the organization staff on the main components of the project and request for the successful involvement in the implementation
3. Detailed analysis of the working flow, setting the technical specification
4. Development, Configuration, Functional testing
5. Functional Integration and Customization
6. Users formation and training
7. Data migration
8. Operational testing
9. Start-up
10. Providing technical support for users
11. Maintenance, warranty
12. Functional system

Our case study shows that the organization decided to implement an Informational System in order of making its public services profitable and more transparent.

4. Study case: The need of an Integrated Informational System Implementation

This study case presents the situation of an education organization confronted with the need of structuring and modernizing its Informational System.

The problem arises at a moment when the public educational system must face up to increased demands on the quality of the services provided as the private sector penetrates the field of education and higher education, which implies introducing in this field the rules and functioning principles specific to the market economy.

Today the ERP systems have become compulsory for the organizations that wish to maintain and improve their acquired position on the market.

The system must offer structured information starting from reflecting the entire activity of the company to the performances on department level, with details on specific activities. A standard pack can be a solution for small and medium businesses, but big companies require systems adjusted to their own needs, able to acquire and manage information from large-scale networks developed in heterogeneous environments.

When an organization decides to develop or implement an Informational System, there are two possible ways of doing it:

- with its own resources;
- by externalizing the IT services: choosing a specialist consultant firm.

Outsourcing the IT infrastructure to data centers leads to increases in the staff productivity in a company, allows rapid resumption of critical operations and protects data from accidental or deliberate loss and damage. Moreover, it reduces and eliminates potentially devastating costs in case of major disruptions.

”Dunărea de Jos” University of Galaţi started by implementing an informatics system by its own resources, planned deadlines for developing and putting into function the modules of receiving study taxes, various education taxes, analytical records of students.

Then a module of cashing in and bank payment as well as that of centralizing situations on faculties and forms of study was developed. The modules for students' grants records, accommodation and transportation were separately developed, by accessing the same database. Apart from these modules there were applications for the students' academic situation which managed a dedicated database.

The financial books applications were difficult to maintain, the database server technology was outdated and the client applications were not compatible with the new versions of the operating system.

The lack of flexibility and interoperability of existing modules and applications led to identifying the need of integrating it into a functional informatics system based on an appropriate infrastructure and a platform which would allow a later extension towards a decisional support system (DSS).

The two possible approaches were analyzed considering the financial and human resources effort, both the direct costs of acquiring and implementation, and the indirect ones related to updating the infrastructure, the application and databases servers, upgrading the network infrastructure, a specialized software for developing the applications, a software for DB management and licenses dimensioned for the necessities of the organization, choosing to setting up an integrated extended ERP system (including CRM package), EERP solution configured and customized to the specific of the activity.

The University Management System - Academis platform is an integrated software solution, with ERP SAP business modules and UMS for customers relationship management, which ensures the efficient administration of all the activities, processes and flows developed in the academic environments and the online-access of the students, as users of applications, to their scholar and financial reports.
According to our case study findings, the Gantt chart for implementing an IIS is detailed bellow:

Table 1: Project Gantt Chart

Stage	Phase	Action	Period
I. Solution selection	I.1	Analysis, evaluation and optimal solution selection	1 month
II. Master plan, Communication, Organization	II.1	Informing the departments managers on the new system and the implementation methodology	2 days
	II.2	Establishing the project team	½ day
	II.3	Establishing the advisers	½ day
	II.4	Establishing the implementation general plan	1 day
	II.5	Establishing the necessary additional computer equipment and software for the operationalization of the new system	½ day
	II.6	Determining the financial resources for project implementation	½ day
III. Analysis	III.1	General analysis	3 days
	III.2	Detailed analysis	10 days
IV. Design, Installation, Functional Testing	IV.1	Technical specification design Development	10 days
	IV.2	Test scenario design	3 day
	IV.3	Functional testing	2 day
V. Integration and Customization		Configuration, parameters setting	2 days
VI. Training and Formation	VI.1	Training of those responsible for implementation actions	2 days
	VI.2	Organizing training courses for top and middle level managers	3 days
	VI.3	Forming of coordinators and advanced users	5 days
	VI.4	Training of users	5 days
VII. Data migration		Data Export / Import	5 days
VIII. Operational testing	VIII.1	Workflow testing	2 days
IX. Start up	IX.1	Functional system Productive work start up	1 day
	IX.2	Acceptance report	1 day
	IX.3	Final reception	1 day
X. Technical support		Assistance and technical support	1 year
XI. Warranty		Maintenance under warranty	2 years
XII. Post-warranty		Updates under post-warranty period	Contractual

Source: own

After successful completing these steps the organization could switch to the next level, setting up a business intelligence system.

The first step to scheduling the project is to determine the tasks that the project requires and the order in which they must be completed. Time is a critical element for the successful finalization of the project. We are using the Program Evaluation and Review Technique (PERT) in order to identify the minimum time needed to complete the implementation project.

Table 2: Project Task T_E determination

Task	Prede-cessors	Type	Length (wks)	To (days)	Tp (days)	Tmp (days)	T_E (days)
I. High Level Analysis		Sequential	1 wks	3	6	5	4,83
II. Hardware Installation	I	Parallel	2 days	1	5	2	2,33
III. Core Modules Analysis	I	Sequential	2 wks	7	15	10	10,33
IV. Design & Developing Core Modules	III	Sequential	3 wks	10	20	15	15,00
V. Installation & Functional Testing	II, IV	Sequential	1 wk	3	6	5	4,83
VI. Integration and Customization	III, V	Sequential	1 wk	4	7	5	5,17
VII. Training and Formation	III	Parallel	2 wks	10	12	10	10,33
VIII. Data migration	IV	Parallel	2 wks	7	15	10	10,33
IX. Operational testing	VI,VII	Sequential	2 wks	7	10	10	9,50
TOTAL Assistance			**14½wks**	**52**	**96**	**72**	
X. Technical support		Sequential	1 year	25	50	28	31,17
XI. Warranty		Sequential	2 years			25	
XII. Post-warranty	XI	Sequential	contractual				

Source: own

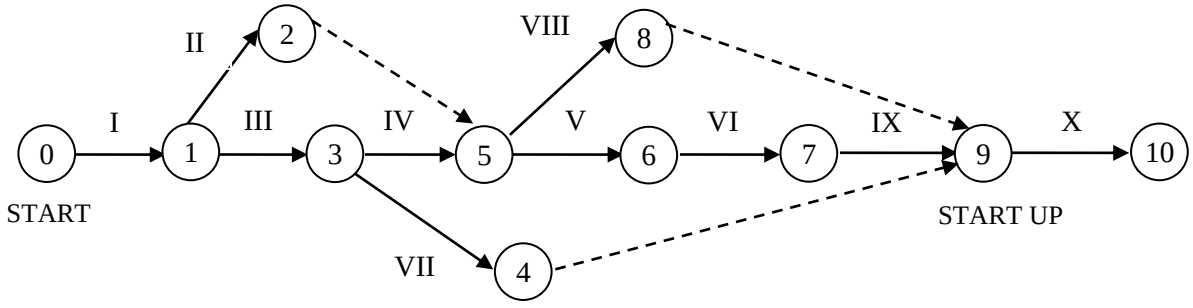
- To *(optimistic time*): the minimum possible time required to accomplish a task (chances are about 1% the activity will be completed within the optimistic time)
- Tmp *(most likely time*): the best estimate of the time required to accomplish a task
- Tp *(pessimistic time)*: the maximum possible time required to accomplish a task, assuming the worst scenario
- T_E (*expected time*): the best estimate of the time required to accomplish a task (the implication being that the expected time is the average time the task would require, if the task were repeated on various implementation experience over many clients).

$$T_E = (To + 4xTmp + Tp) / 6 \qquad (1)$$

Crashing Critical Path

The critical path is the longest path of planned activities to the end of the project. Using the CPM technique to determine the path times, we add the task durations for all available paths drown in Figure 2 and calculate the path that takes the longest to complete.

Figure 2: PERT Diagram of the IIS Project



Source: own

The critical path is represented by the activities 0-1-3-5-6-7-9 and the duration of this pathway is equal to 50 work days, i.e. two months and a half.

The objective of this analysis is to accelerate the implementation process in order to reduce costs and time. The expected late finish time for the project is LF = 40 work days, the equivalent of two months.

The time required to execute the task can be reduced for an additional cost. The conclusion of the decision-making process based on the analyses of the PERT diagram of the project is that there are two areas that need to be accelerated. The IIIrd and IVth activities were identified as time-consuming and potentially critical for overcoming the project deadline. Close monitoring and analysis of project progress was the method for cutting off the delays. The key factor for reducing the elapsed time for the detailed analyses from 10 to 5 days and the technical specification design from 15 to only 10 days was the decision of bringing the project team members working together in a collaborative environment.

This platform based on free web-portal with forum and discussion facilities was provided for the project management and should allow to all users, inside and outside of the organization, to build relationships and business activities by intranets and external communication channels. The key words of using these interactive tools of the web 2.0 are sharing and exchange information. The using of social computing applications is opening new ways to work: bottom up, open, interconnected, distributed, stimulating the professional and the users to participate actively in this project. (Consoli, D., 2011; 2012)

The goal was to improve the quality of the support services and technical assistance provided, establishing a closer collaboration between the users and the software services provider, for monitoring and prompt solving of the reported dysfunctions. The solution chosen was to implement a demand management system for bug and issue tracking, integrated with other software project management and mail tracking applications. The software debugging system to be implemented is Web-based and allows for ticketing and automatic notification of problem reception, with the demand's recorded date and time, priority degree allotted, estimated time for problem solving, status modifications being notified along the entire life cycle of the problem. This system complete the technical support ensured through the online support Help Desk to solve urgent problems.

The advantage of this solution is to ensure a shorter response time and effectively to any problem raised by the software user. Automatic referral management will reduce the response time to a ”Guaranteed intervention time” and, respectively, ”Guaranteed solving time” (Time elapsed from a demand registration until it is solved must not be longer than the guaranteed solving time according to the demand's priority degree). Centralized records of all the referrals will allow for a correct management of the new demands and the unsolved or just partially solved demands in the project, according to their status.

IT to BT - if you change that one word from *information* technology to *business* technology, you begin to change the way IT people work and the way they think about their jobs. (Wailgum, T., 2008)

By changing that one word, you’d get the mentality of the people working in technology to change how they think and ultimately to create different behavior that is much more business-oriented. (Moore, C., 2009)

The expected results of an IIS Implementation are the technology embedded advantages:

1. Flexibility allows for adaptations and updates in order to cover prospective necessities
   The flexibility principle is a basic principle in designing technology because of market evolution, high mobility of demands, both quantitatively and qualitatively, which requires an increased capacity of adaptation, of rapid adjustment of technology. Flexibility poses new fundamental problems to such a system right from the design phase: adaptability, integrability, dynamic conception to manage the complexity and allow the system optimization
2. Scalabilities, the possibility of further extending the number of licenses according to the evolution of the activity
3. Reliability throughout the life cycle: Reliability is defined in the design stage, made in the implementation phase, provided by control and maintained during the operational phase
4. Security design aims to address the following issues: reliability, maintainability and protection against failure in normal, predictable and also in unexpected or less predictable circumstances
5. Maintenance, as the complexity of the systems and the maintenance costs and the guarantee of the current operating system increase, special attention is given in the design to maintainability, so that bugs can be rectified in a short time

6. Ergonomic features, consisting in providing a friendly interface, facilitating routine work, anticipating errors, validating intermediary operating steps allowing corrections
7. Integration between the workflows

5. The key factors for a successful implementation:

5.1. The involvement of the top management

The implementation decision should be promoted and sustained by the management of the organization. Such an important project which involves the implication of all the departments can't be unrolled successfully without the sustained intervention of the board.

Strong management involvement in the development of implementing activities determines the organization employees' involvement through sessions of project stages analysis, levels training of the staff, incentives for deadline respect.

The role of the project team in the implementation activity coordination: As it is an integrated system that engages all departments concerned, to prove its efficiency each user of the system should be made aware that each must fulfill their responsibilities and correlate their activities with those of the other users.

The involvement of departments have the effect of positive answers to management initiatives and thereby increase employee performance, leading to increased flexibility and a long term strategic innovations of the company.

Management Involvement => Staff involvement => Long-term strategic innovation

5.2. Change Management

Anticipating and managing the major changes that occur within the organization as a result of the IS implementation: This project type has an impact on the organization as a whole, requiring rethinking and redesign of organizational mechanisms, reorganizing the internal flows and adapting to the needs of the informational system.

5.3. Pilot Phase

Creating a pilot project representative for the activity of the organization, detailed testing of the procedures in the pilot phase, the results of this stage being particularly useful in further expansion of the project in other areas of activity.

The operational testing of the application in a pilot implementation is recommended, that will initially involve a small area of activity, allowing limitation of the implementation effort to the pilot field chosen and minimizing the impact in case of failure.

The isolation of the Pilot application of the other functional applications ensures the opportunity to remedy the errors, to identify the necessary adaptations and resume phases of that stage.

5.4. Compliance with the methodology and implementation plan

Thorough documentation of every phase and stage of the project is a sub-project in itself, which aims to improve the management of the implementation project and to follow the evolution of the project, according to the Capability Maturity Model Integration. (Paulk, M., 2002)

Compliance with the methodology throughout the project is one of the project manager's duties, together with regular checking of the quality of documentation: deliverables of each phase synthesis, testing scenarios, configuration, training and formation reports, workflows, requests for modification or customization of the application, update and version management, protocol of partial and final receiving, audit reports, etc.

Taking into consideration the implementation duration, evolution of the project team in time, the tendency to make changes including in the methodology of implementation, complying with the targets originally set is a prerequisite to framing the project in time and within budget. Each step must be completed by a synthesis deliverable report.

5.5. Project Phasing

We must break the overall project into several mini projects, to not completely disrupt activity. ERP projects that have succeeded have often made the implementation over a long period of time and broke the project into several stages. Many bring up a sample of users, especially users who agree with the added value of such a tool. In this way, they could dive into the project first and see the pitfalls to be

corrected. When a consensus is passed successfully, the test group of users could help and train employees in other departments and help their colleagues with particular situations. (Conso, P., 2006)

5.6. Setting objectives, responsibilities and tasks

Every project needs an internal champion, calls as the "Super User". Projects that are successful had one. It must be a person from the administration that is recognized and appreciated by all, maintaining team cohesion throughout the project and who should get all the advanced training necessary and should know everything about the system.

Distribution by groups of staff roles enables balanced distribution of responsibilities and a better security in terms of access to different functions of the application, providing an overview of the occupancy of each user and a means of anticipating busy periods loaded with more activity and to plan additional resource requirements. He is also responsible for the accuracy of the database, focusing on the reliability and safety of the data.

In fact the super user is essential to the success of ERP adoption by the team, monitoring of completed activities and of those left to be performed, by permanent consultation with experts in each field of activity.

5.7. Technology transfer and innovation promotion

The process through which a system of scientific and technical knowledge on processes, work procedures consisting of documentation, workflows, know-how, acquired experience used for the development, operation, maintenance of the application, namely for the development of certain categories of activities is integrated into a context, into the implemented new technology system. (Bogdan, I., 2006)

The reverse transfer, the feedback reaction is the one that stimulates the evolution of the information system in the direction of increasing the interdependent efficiency of the system components.

The technology importer is not just a receiver, he has knowledge in his field of activity, knowledge which, associated to that of the technology provider, produces a synergistic effect, the success being conditioned by the size of the enterprise which supplies the technology compared to the size of the beneficiary and the degree of the partners' mutual commitment. (Nicolescu, O.; 2004)

5.8. Implementing a risk management system

In our opinion, the main risks that can be quickly identified are:

Table 3: Project Associated Risks

	Identified risks (IR=Internal risk, ER= External risk)	**Risk type**
i.	Change of the project team	IR+ER
ii.	Modification or redistribution of tasks within a department	IR
iii.	Project changes	IR
iv.	Design omissions and errors	ER
v.	Legislative changes that occur during the implementation	ER
vi.	Failure to comply with deadlines for each stage (for example: to collect admission fees it was necessary to find a replacement solution)	ER+IR
vii.	Failure to achieve quality standards established by the quality plan (examples: program bugs left unsolved, reports failing to meet the deadlines or incomplete reports that accompany the completion of each phase or stage)	ER
viii.	the training, documentation, task procedures delivered in time diminish the operating errors which take extra time for checking and correction	ER+IR
ix.	poor communication within the project, communication deficiencies	ER+IR
x.	Lack of a detailed planning of the project and poor implementation monitoring	ER+IR

Source: own

The migration and check of the transferring data stage must not exceed the planned schedule. Failure in the duration of this stage generates the risk of repeating this stage at a later date, restarting implementation, for example the date of the implementation start up was set for Jan, 1st, and must be delayed to March, 1st, which requires resuming all the activities of this stage, preparing a new set of data exported from the initial application to the newly established date, running and checking the results of data imported script with the efforts of all involved.

After finishing this stage can move on to the parallel running stage for a maximum 3 months recommended period.

On the course of this stage the new application is tested, the reports and the results are checked and configured with the ones emitted in the old application.

After successfully finishing this stage it is possible to give up working in the old application, only after testing a closing month and validating it in the new application, associated with issuing the reports (analytical balance, balance sheet, cash book, bank registry, etc.)

Premature separation from the old application without checking the new application reports on data entered manually or taken over by migration carry the risk of having no continuity in any of the applications and the impossibility of submitting reports on time.

Extended work in parallel in the application that users were accustomed to working with, after the start up of the new application, represents a further effort of operating the data which is not justified as long as the new application was validated.

Some activities are different from the work in the old application, e.g. registration of an Inventory Object or a Fixed Asset acquisition is done through a sequence of steps:
Approved report / Ordering / IRN (input reception note) / IO.

Failure to comply with these standards of quality or the implementation period may lead to the exceeding of costs.

6. Conclusions

Promoting technical progress can be done only on a very clear vision on the organization and its objectives. Such a vision must include the following features: planning, guidance, motivation and management per se.

Formulating a vision for promoting technical progress by the proposed investment is done in three main phases: analysis and diagnosis, design and implementation.

The analysis and diagnosis phase involves researching the existing situation, answering to the following questions: how to work and what are the new techniques, which are the strengths and weaknesses of the system, both in terms of technology and manpower: who is the competition, what it is characterized by and how it can be overcome, which are the concrete objectives, the management methods, who are the current and the potential customers.

The design phase results from the responses to the previous phase and the success of the decision making process depends largely on the number and quality of the alternatives taken into account.

The communication phase is also essential by setting-up and using of communication channels, mail groups: operational and correspondingly, top management information.

Simulations of the main workflow of an organization on test cases, analysis and complex statistical models are some of the amenities offered by a modern information system, needed by a company to face the new economic climate.

For the successful implementation of our case study an important factor was the use of technological platform that ensures collaborative tools for content management, such as: forums, chat, bug tracking and wikis for project documentation, work procedures and scenarios. Accessing information faster helps finishing projects on time and within budget.

Business-Critical Solutions like ECM/ECP (Enterprise Content Management/Enterprise Collaborative Platform) speeds project completion with centralized content management.

According to our conclusions, good knowledge of technology has constituted an advantage, implementing a quality management system, the basic condition for the cooperation, ensured the continuation of the activity of the project in the process of implementation.

The effect of redesigning the business process depends on how well it is implemented, coordinated and monitored.

On the other hand, equally important are development strategies, decision-making processes the organizations rely on.

In the current economic environment, characterized by competition and the rapid evolution of requirements, optimizing the decision-making process depends on the speed of adaptation of information systems to follow this evolution so as to be able to provide real time information, enabling timely decision making, competitiveness of provided services in terms of increased demands on the IT services market and the innovative nature of products offered in line with the evolution of technology is a priority goal for an IT company.

Experience gained while implementing Enterprise Resource Planning systems shows that in-depth analysis of internal and external factors that influence the company's evolution is critical for the transformation of such a system into an effective tool available to managers.

7. References

- Bogdan, I., (coord.), *Elemente de metodică a cunoaşterii ştiinţifice şi valorificării cunoştinţelor*, Editura Universitară, Bucureşti, 2006
- Conso, P.; Hemici, F., *L'entreprise en 20 Leçons*, Ed. Dunod, 2006
- Consoli, D., 2012, The adoption and use of ERP Systems in Small-Medium and Large Companies, Hradec Economic Days
- Consoli, D., 2012, *The Integrated Software Platform in the Implementation of an Enterprise 2.0 Model*, ICT for Competitiveness 2012 Conference Proceedings
- Hamerman, P.D., *Trends 2011: ERP Customers demand better flexibility, cost transparency and mobility, An Interview with Andrew Magarie*, ForresterResearch (Jan 11, 2011) [Online]
- Jacobs, F.R; Bendoly, E, 2003. *Enterprise resource planning: Developments and directions for operations management research*, European Journal of Operational Research, 2003, Vol. 146, No.2
- Jeandet, T., *Business Information-Democracy Solutions*, Business Intelligence Day Conference, Bucuresti, 2004
- Martens, C.; Hamerman, P.D., *For Business Process Professionals: The State Of ERP In 2011: Customers Have More Options In Spite Of Market Consolidation, An Interview with Andrew Magarie*, ForresterResearch (May 02, 2011) [Online]
- Moore, C. *From Information Technology To Business Technology: An Interview with George F. Colony and Charles Coit*, ForresterResearch (September 16, 2009) [Online]
- Nicolescu, O.; Plumb, I.; Pricop, M.; Vasilescu, I.; Verboncu, I. (coord), *Abordari moderne in managementul si economia organizatiei, Editura Economica*, 2004, Vol. 1, p.161; Vol.4, p.190
- Parker, B., *Your Business Has Changed - Is it time to change your ERP?*, CIO Magazine (2011) [Online]
- Parker, B. *Operations Management White Paper: The Other Enterprise Platform*, IDC 2011. Available at *www.idc.com*
- Paulk, M., 2002, *Capability Maturity Model for Software. Encyclopedia of Software Engineering*
- Ried, S., *Market Overview: The Middleware Software Market, 2009, Software AG And TIBCO Challenge IBM, Microsoft, Oracle, And SAP, An Interview with Jonathan Penn, Miroslaw Lisserman*, ForresterResearch (Dec 30, 2009) [Online]
- Wailgum, T. *Why ERP Systems Are More Important Than Ever?*, CIO Magazine (2008) [Online]
- Wailgum, T. *You say IT, Forrester says BT: What's the Difference? An interview with George F. Colony*, CIO Magazine (2009) [Online]
- *http://www.compareerpsoftware.com/vendor-review/dolibarr-erp-software-review/*
- *http://www.erpuniversitar.ase.ro/*

INVESTMENT DECISION ANALYSIS IN INFORMATION SECURITY

GHEORGHE Mirela,
Professor Ph.D., Faculty Accounting and Management Information Systems of The Bucharest University of Economic Studies, Bucharest, Romania, mirelaghe@gmail.com

Abstract: *The purpose of this paper is that of presenting the economic indicators specific to an investment decision analysis in information security. The motivation for this research originates from the need of having some economic models that will allow the person in charge of making decisions in the field of information security to choose the best technical solutions with optimum investments. It is very hard to evaluate and measure the benefits generated by the security solutions in the field of IT. The paper will explore the facilities offered by the economic indicators specific to the field, such as the return on investment (ROI/ROSI), the net present value (NPV), the internal rate of return (IRR), the dispersion, the standard deviation and the variation coefficient with the purpose of helping the decision makers accomplish optimum investments for reducing the security risk to an acceptable level.*

Key words: Investment Decision, Information Security, Return On Investment In Information Security, Net Present Value, Internal Rate Of Return.

JEL classification: M15

1. Introduction

In the last decade, IT integration turned into a "sine qua non" condition for the business environment dictated by the need to guarantee real time information processing and transmission, Internet connectivity enhancement and implied productivity growth, the need to ensure competitive advantage in an economy ever more dynamic and complex, specific to today's society, a society based on knowledge.

The evolution of the present information environment is a natural one to which the organizational environment must adapt, integrate the best techniques and instruments which can grant it transparency and the data necessary to determine, for example, which are the priorities in accomplishing investments so that these support the organization's objectives and bring an added value. The opportunities generated by the increasingly broader IT integration are undeniable, but on the other hand we must admit the fact that at organizational level there have appeared new risks specific to the informational environments (digital business) that can lead to substantial losses if ignored. In the digital era, businesses must rely on a proactive and systematic frame, promote change management and be ready for any disaster scenarios.

Data and information security, elements that constitute the valuable assets (intellectual capital) in the knowledge based society, is essential for the competitiveness and sustainability of an organization. Cybernetic attacks have become more and more complex. The world economic crisis is one of the factors that determined the increase of the level of cybernetic crime at global scale right since the year 2008. The high number of viruses, worms and other malicious codes (malware), phishing attacks, botnets networks expansion and the worrying growth of spamming confirm the severity of the problem. Facing these threats, investments made in ensuring information security become a pressing problem for the manager of any entity. Often they are confronted with questions to which the answers are not yet known:

- How will the lack of information security affect the organizational environment?
- Which will be the impact of an information security breach over the organizational environment?
- Which is the optimum level of investments to ensure an adequate security?
- How will the company know if the investment decision made based on the feasibility studies offers the most adequate security solution?
- How will the benefits generated by investments in information security be evaluated?

The central objective of this paper is to present economic indicators specific to an analysis of the investment decision in information security. To achieve this goal different security risk reducing scenarios have been analyzed from the point of view of the financial indicators: investment return, presented net value and internal return rate, but also from the point of view of the statistical indicators: the dispersion, the standard deviation and the variation coefficient.

The structure of this paper is the following: chapter 1 performs an analysis of the literature concerning the concept of information security risk. Chapter 2 describes the content of the security risk

management process, underlining risk assessment. Chapter 3 offers a presentation of the main economic indicators for evaluating decisions in security information investments. In chapter 4 we detail an investment decision from the perspective of three scenarios, using field specific indicators such as ROI/ROSI, NPV and IRR as well as dispersion, standard deviation and the variation coefficient. The last section, conclusions, offers some observations and limits over the executed research.

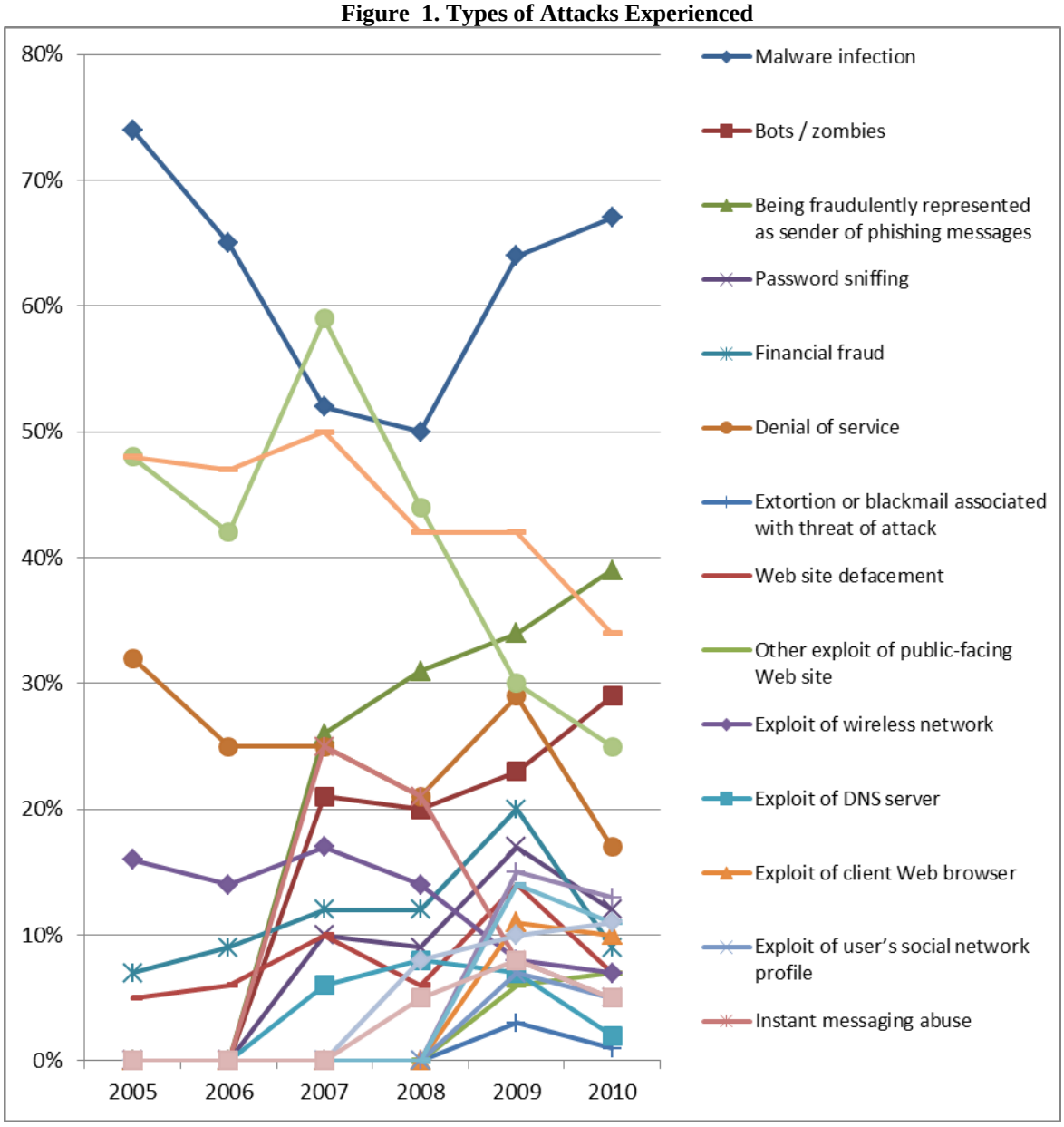
2. Information Security Risk

Any security strategy must take into account three fundamental elements of an informational system's security, which take the form of its goals: *confidentiality* (the ability of protecting information sensible from unauthorized persons), *integrity* (ensuring accuracy, validity and completeness of information, by protecting against corrupting it or degrading) and *availability* (ensuring prompt access to operational information). These three elements, known under the name of *CIA Triad*, form the ground on which electronic information security is based on, being widely recognized as basic principles in adopting a model in this sense. In time, though, other specific security strategy elements took shape and have been accepted as being part of upright model. Thus, at the present moment, the following model is considered fit, a security model characterized by the three elements that make the CIA Triad, to which are added others as:

- *Authentication* (ensuring the authenticity of a person, organization or device that generates information inside the system);
- *Authorization* (granting specific privileges to users of the informational system through administrative tools, monitoring and security control);
- *Access control* (granting access to the information only to authorized personnel)
- *Responsibility* (the possibility to identify users that are responsible of generating or modifying information);
- *Safety* (trust in truthfulness of the information in the system);
- *Non-repudiation* (no action that affects information security can be denied at a future moment) (Keifer et al., 2004).

Information security assumes ensuring data and information protection against threats that could exploit the system's vulnerabilities. Actually, any information, assimilated with the organization's asset, is exposed to inside or outside the organization threats (IT system breakdown, hacker intrusion, human error, viruses, financial frauds, and so on). Even though in the last 10 years, the evolution of the IT security technology has recorded a significant progress, the level of security of the IT networks and systems hasn't been noticeably enhanced (Schneier, 2004; Whitman, 2003). Cybernetic attacks have become more diversified. The great number of viruses, worms and other malicious coding (malware), phishing attacks, botnets networks expansion and the worrying growth of spamming confirms the severity of the problem (CSI, 2010) in figure 1.

Figure 1. Types of Attacks Experienced



Source: (CSI, 2010)

Turban (Turban, 1996) defines the risk as being "the probability that a threat materializes". When talking about informational systems, the **risk** is a function that depends on ***threats*** (those events that cause losses), ***the probability*** of a threat fulfillment, IT system ***vulnerabilities*** (the weaknesses of a system) and the ***value of exposed assets*** (information that must be protected):

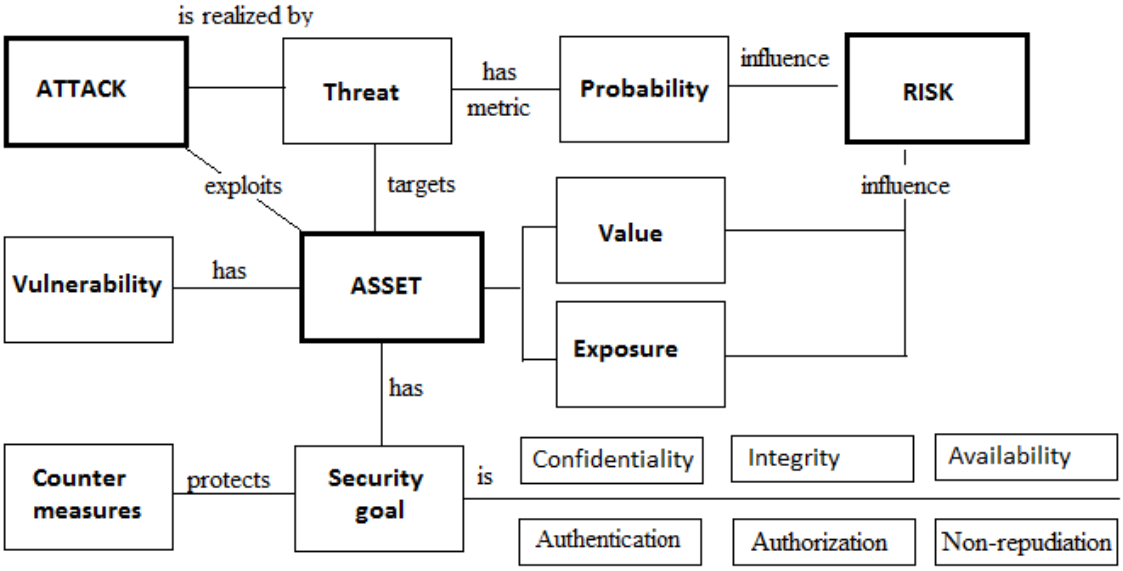
Risk=f (prob, Threats, Vulnerabilities, Value of assets)

The definitions found in literature associated to security risk are multiple. Just as an example we list the definition provided by NIST (NIST, 2002) "**Risk** is a function of the ***likelihood*** of a given threat - source's exercising a particular potential vulnerability, and the resulting ***impact*** of that adverse event on the organization." Derived from the standard definition of risk presented earlier, the typical risk formula is accepted as follows:

Risk= Impact * Likelihood

Building from these definitions, figure 2 puts in correspondence the relations that are established between the components of information security risk.

Figure 2. The components of information security risk



Source: (Miede, 2010)

3. The Management Of Information Security Risk

The management of information security risk is a global process that integrates the identification and the risk analysis to which the organization is exposed, that evaluates the potential impact over the organization, and which decides what actions must be taken for eliminating or reducing it to a certain acceptable level. Talking just for the case of information security management, we can say that the process requires an identification and the evaluation of the organization's informational assets, of the consequences of security occurrences, of costs and benefits generated by the investments in informational security. The standards and guides that regulate this field (ISO 27000 series, NIST, and so on) include in the process of security risk management the following steps:

- Security risks identification
  - Identifying and classifying the organization's assets that must be protected
  - Identifying the threats brought upon the informational system
  - Identifying the system's vulnerabilities
- Evaluating security risks
  - Estimating the threat's probability of fulfillment
  - Estimating the impact upon assets
- Choosing the risk handling strategy
  - *Risk avoidance* – eliminating the uncertainty by not initiating the activity rated as being risky.
  - *Risk transfer (risk outsourcing)* – consists in assigning the property upon a risk to a third party (through bonds, insurances, contractual clauses).
  - *Risk decrease* – reducing the risk occurrence probability or its impact, through the implementation of some additional control measures.
  - *Risk acceptance* – this type of risk response applies to those inherent risks for which the exposure is lower than the risk tolerance.
- Cost-benefit analysis for the security solutions.
- Monitoring the efficiency of the implemented controls.

Evaluating the security risk

Practically speaking, there are two approaches in risk level quantification: a quantitative approach and a qualitative one. The quantitative approach is based on mathematical reasoning providing numerical assessments of the security risk, while the qualitative analysis offers relative assessments of the same risk, the method itself being based on brainstorming techniques and questionnaires. Both approaches, quantitative and qualitative have their specific advantages and disadvantages.

Quantitative analysis

The quantitative analysis represents an approach which is based on some calculus and formulas specific from the field of statistical mathematics. Two basic formulas can be found in the literature, related to the quantitative risk analysis: single loss exposure (SLE) and annual loss expectancy (ALE). These are calculated as follows:

SLE = Asset Value (AV) * Exposure Factor (EF) (1)
ALE = SLE * Annual Rate of Occurrence (ARO) (2)

The single loss exposure represents the expected loss related to a single occurrence of an event, and the prediction of the value of the damage that can be inflicted as consequence to this event can be used as a solid ground when choosing the security strategy. Many security risk estimation techniques calculate this indicator based on a formula that incorporates the exposure factor and the asset's value.

A simplified example, proposed by Rok Bojanc and Borka Jerman Blažič (R. Bojanc & B. J. Blažič, 2008), explains this approach for the next situation: a WEB data server with a value of 30 000 euros is infected with viruses, fact which leads to a loss of 25% of its value. In this case the loss related to this unwanted event will be:

SLE = AV * EF SLE = 30000 * 25%=7.500 euros.

Annual loss expectancy is used because of the fact that in the field of risks one cannot operate with certitudes, but only with probabilities. For the above example, it is appreciated that ARO = 0.5, which means that this type of event (virus infestation) can occur once in 2 years. ALE= 7 500 * 0.5 = 3750 euro.

Qualitative analysis

While quantitative analysis is based on complex formulas and monetary values of the variables, qualitative analysis is based on subjective reasoning. The fundamental elements are the same: asset value, threat frequency, impact and efficiency of preventing measures, but these elements are evaluated in subjective terms, such as "high level", "medium level" and "low level" (Krause & Titon, 2002). Even if qualitative variables aren't expressed under numerical form, they can still be set in order (High>Medium>Low).

4. The Decision Of Investing In Information Security

Risk decreasing through investments is the main strategy of risk management. The purpose of these investments is of decreasing risks associated with informational security According to a survey executed by CSI (CSI, 2010), the average of investments in information security is 3-5% of the IT budget and less than 1% of the IT security budget allocated for employees training. One of the reasons for this situation is the lack of trust in the existing security models, and a second reason would be represented by the fact that most of the organizations look at expenses accomplished in the field of information security as simple expenses and not as investments.

Gordon & Loeb (Gordon & Loeb, 2002) stated that: "In order to protect confidentiality, integrity and availability of information, ensuring at the same time authenticity and non- repudiation, organizations invest large amounts of money. Because investments in security are concurrent with other investments it is not surprising that CFO requires a rational economic approach of these expenses". In 2002 the same researchers proposed a simple and yet general model for an optimum level in the field of information security. The Gordon & Loeb model is based on three parameters: λ – which represents the monetary loss that a company would record in the moment of the occurrence of a security breach, t – the threat fulfillment probability and υ – vulnerability. Judging on this model, the specialists have demonstrated that the optimum level of investments in information security cannot exceed 36.8% ($\approx 1/e$) of the total value of informational assets (expected loss due to a security breach). Reactions of the field's specialists didn't hesitate to appear; there are authors that opposed this model, demonstrating that there can exist situations where the optimum investment can exceed 50% of the asset value (Willemson, 2006), but there are some other authors that have expanded this model (K. Matsurra, 2008) by introduction of the concept of "productivity space of information security".

Classical indicators for cost and benefits quantification for the investments in information security rely on the return on investment (ROI), net present value (NPV) and internal rate of return (IRR)

or a combination of the above. Also, this analysis can be executed based also by computing dispersion, standard deviation and the coefficient of variation.

The costs of investments in information security include usually the cost of the technical equipment, software, theirs installation cost, but also additional costs such as maintenance costs, updating, employees training. On the other hand, the benefits of investments are the equivalent of losses that could record if the information security risks would materialize.

Return on Investment – ROI

ROI can help an organization decide which of the possible options offer the greatest value for invested cash. ROI is equal to the actual value of the accumulated net benefits over a certain period of time, divided by the initial investment costs.

$$ROI = \frac{\text{Benefits} - \text{Cost of investment}}{\text{Cost of investment}} \quad (3)$$

In the field of IT it is very hard to evaluate and measure the benefits. For example solutions for security such as firewall systems, IDS, antivirus software cannot generate benefits that can be measured. In this context, benefits can be measured as the difference between ALE without investments and ALE with investments (Gordon & Loeb, 2006).

Benefits = ALE without investments – ALE with investments (4)

$$ROI = \frac{\text{ALE without investment} - \text{ALE with investments} - \text{Cost of investment}}{\text{Cost of investment}} \quad (5)$$

A practical model for computing the return on investment in information security, proposed by Sonnenreich (W. Sonnenreich, 2006) and adopted by ISACA (ISACA, 2010) through the guide G41 "Return On Security Investment (ROSI)", introduces the **ROSI** metric (Return on Security Investment):

$$ROSI = \frac{(\text{Risk exposure} * \%\,\text{Risk mitigated}) - \text{Cost of security investment}}{\text{Cost of security investment}} \quad (6)$$

Although it is a parameter very often used in practice, ROI/ROSI is not recommended for use when the investment is distributed over a large time span (>1 year) because it doesn't allow updating the monetary value according to an annual depreciation rate.

Net Present Value – NPV

NPV can be defined as being the difference between the sum of financial input flow, updated with an adequate cost of capital (update rate), and the present value of the initial investment.

$$NPV = \sum_{i=1}^{n} \frac{CF_i}{(1+rate)^i} - I \quad (7)$$

where n – number of periods, CF – cash flow, rate – update rate, I – investment cost.

The net investment value can be considered as being the measure of the surplus which the investment obtains in excess to the necessary income or an error margin relative to the size of investment.

Analyzing an investment of 50.000 euros made in a security solution to prevent a loss of 75.000 euros in the first year of solution deployment or in the 3rd year after the deployment, at an update rate of 5%, although the return on investment has the same value of 50%, NPV will be different in the two situations (NPV = 21428,57 € and NPV = 14787,82 €). The NPV indicator takes into account the time moment when the benefit was generated (in our case it resembles with the prevented loss) offering the decision maker a real expression of the investment value.

Internal rate of return – IRR

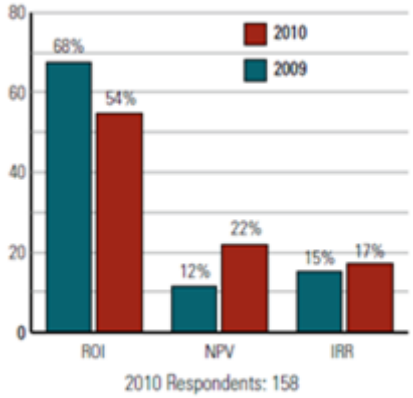
IRR is that rate which will make the net present value drop to zero. It is the income generated inside the company that will be acquired its lifespan.

$$\sum_{i=1}^{n} \frac{CF_i}{(1+IRR)^i} - I = 0 \quad (8)$$

Each of these financial measures has their specific advantages and drawbacks. ROI is destined for use when evaluating investments done in the past, while NPV and IRR are used when discussing a new potential investment (Gordon & Loeb, 2003). In most of the cases, NPV and IRR are better indicators than a simple ROI (Gordon & Richardson, 2004).

According to the CSI study in 2010, 54% of the organizations use ROI, 22% use NPV and 17% use IRR (figure 3).

Figure 3. Percentage of using ROI, NPV and IRR indicators in organizations



source: (CSI, 2010)

5. Case Study – Investment Decision Analysis

A corporation decides making some investments for reducing the security risk (example adapted from Rok Bojanc & Borka Jerman-Blazic (2008)).

After a risk analysis, it is estimated that the annual loss due to certain breaches of security is of 3 000 000 euros. The corporation, at present time, has specific controls implemented that cut the security risk by 80%. The goal of the organization is to choose the most adequate solution for reducing the security risk to at least 10%. It is predicted that the investment will be made over a period of 5 years, and the corporation's managers wish to analyze the investment decision starting from 3 scenarios, considering an update rate of 5%.

In this way, the first scenario is represented by a less expensive solution developed in-house of a security system which will allow reducing risk at 10%. The investment cost is of 100 000 euros and each year maintenance expenses of 15 000 euros are predicted.

The second scenario is represented by a solution offered by a specialized firm which will allow reducing the risk to 1%. Acquisition price is of 150 000 euros and the annual maintenance cost is of 5 000 euros, update to the proposed solution of 30 000 euros. Also, the specialized firm will organize in the first year of implementation a course for training the employees valued at 25 000 euros.

The third scenario is represented by outsourcing the information security service. The company offers a security risk reduction to 5%. The cost of the solution is 200 000 euros and another 25 000 for annual maintenance costs.

Which is the best investment decision judging from the economics point of view?

The benefits associated to the 3 scenarios are calculated starting from the loss that could materialize weighted with the difference between the diminished risk predicted and the actual diminished risk:

Benefit $_1$ = 3 000 000*(90%-80%) = 300000 euro
Benefit $_2$ = 3 000 000*(99%-80%) = 570000 euro
Benefit $_3$ = 3 000 000*(95%-80%) = 450000 euro

We will analyze the three financial indicators, ROI, NPV and IRR for the proposed solutions.

$ROSI_1$= (3 000 000*(90%-80%) – 1000000)/1000000=200%
$ROSI_2$= (3 000 000*(99%-80%) – 1500000)/1500000=280%
$ROSI_3$= (3 000 000*(95%-80%) – 2000000)/2000000=125%

Table 1. Cost and benefits analysis for the 3 scenarios analyzed over a 5 years period

		Scenario 1		Scenario 2			Scenario 3	
Year	**Update Rate**	**Benefits**	**Investment cost**	**Benefits**	**Investment cost**	**Additional Cost**	**Benefits**	**Investment cost**
0	5%		1000000		1500000			2000000

1	5%	300000	15000	570000	30000	30000	450000	25000
2	5%	300000	15000	570000	30000	5000	450000	25000
3	5%	300000	15000	570000	30000	5000	450000	25000
4	5%	300000	15000	570000	30000	5000	450000	25000
5	5%	300000	15000	570000	30000	5000	450000	25000

Table 2. Calculation of IRR, NPV and ROSI indicators for the 3 scenarios

	Scenario 1	**Scenario 2**	**Scenario 3**
Initial investment	1000000	1500000	2000000
Cash Flow 1	285000	510000	425000
Cash Flow 2	285000	535000	425000
Cash Flow 3	285000	535000	425000
Cash Flow 4	285000	535000	425000
Cash Flow 5	285000	535000	425000
IRR	285%	344%	212%
NPV	1.133.900,85	2.142.460,49	1.640.027,59
ROSI	200%	280%	125%

Analyzing the results, according to table 2, the profitable solution is given by scenario 2.

Next the three statistical indicators dispersion σ^2(NPV), standard deviation σ(NPV) and variation coefficient C_v will be analyzed.

Tabel 3. Analysis of the 3 scenarios using dispersion, standard deviation and coefficient of variation

	Update Rate		**5%**						
	Scenario 1			**Scenario 2**			**Scenario 3**		
Year	**Probab**	**CF**	**NPV$_i$**	**Probab**	**CF**	**NPV$_i$**	**Probab**	**CF**	**NPV$_i$**
1	10%	285000	171428.57	19%	510000	335714.29	15%	425000	204761.90
2	10%	285000	429931.97	19%	535000	820975.06	15%	425000	590249.43
3	10%	285000	676125.69	19%	535000	1283128.17	15%	425000	957380.41
4	10%	285000	910595.89	19%	535000	1723274.00	15%	425000	1307028.96
5	10%	285000	1133900.85	19%	535000	2142460.49	15%	425000	1640027.59
	NPV		332.198,30			1.198.054,88			704.917,24
	σ^2 (NPV)		1.131E+11			3.91555E+11			2.34602E+11
	σ (NPV)		336296,83			625743,44			484357,04
	C_v		1,01			0,52			0,69

The calculus made shows that the most profitable investment is the one provided by scenario 2 (the highest average NPV) and at the same time the least risky (it shows the lowest standard deviation and variation coefficient).

We observe that the two analysis methods confirm the same scenario as the profitable investment, fact that validates the indicators used during the analysis.

6. Conclusions

In an economy in which information technology has imposed over all fields of activity, the risk topic, more precisely the information security risk, gains special connotations and imposes changes in the fundamental principles of the risk management process. Thus, one can notice that the actual trend is that of promoting a *proactive management* which can allow identifying potential threats, before they materialize and produce adverse consequences among the established objectives. Risk reduction through investments is the main strategy of this process, and the investment decision analysis remains a challenge for any manager.

The paper centered on the analysis of decision making regarding investments in information security, presented in a holistic approach, starting from the concept of information security risk, analyses the information security management process at organizational level, the accent falling on risk evaluation and ends with a description and exemplifying of the principal analysis indicator regarding investments in information security. With this in mind, based on three scenarios that allow security risk diminishing at the entity level, the research offers an economics model of investment analysis using financial indicators such as ROSI, NPV, IRR and statistical indicators such as dispersion, standard deviation and variation

coefficient. The results confirm the same investment solution, a fact that validates the adopted analysis indicators.

7. Acknowledgements

This work was co-financed from the European Social Fund through Sectorial Operational Programmer Human Resources Development 2007-2013, project number POSDRU/89/1.5/S/56287 „Postdoctoral research programs at the forefront of excellence in Information Society technologies and developing products and innovative processes", partner Bucharest Academy of Economic Studies – Research Center for "Analysis and Regional Policies.

8. References:

- CSI (2010), *2010 / 2011 CSI Computer Crime and Security Survey,* available on-line at www.gocsi.com.
- Gordon, L., Loeb, M. (2002). „Return on information security investments: Myths vs. Realities". *Journal of Strategic Finance 84.*
- Gordon, A. L., & Richardson, R. (2004). „The new economics of information security", *Information Week.*
- Gordon, A. L., Loeb, P. M., & Sohail, T. (2003). „A framework for using insurance for cyber-risk management", *ACM, 46(3).*
- Gordon, A. L., & Loeb, P. M. (2006). „Managing cyber security resources: A cost–benefit analysis", *McGraw Hill.*
- Krause, M. & Tipton, H. F. (2002). *Handbook of Information Security Management,* CRC Press LLC, ISBN: 0849399475.
- ISACA (2010), G41 *Return on Security Investment (ROSI),* available on-line at www.isaca.org.
- Matsuura, K. (2008), „Productivity Space of Information Security in an Extension of the Gordon Loeb's Investment Model",
- Miede, A. (2010), „A Generic Metamodel for IT Security – Attack Modeling for Distributed Systems", Germany.
- NIST (2006). *Information Security Handbook: A Guide for Managers*, National Institute of Standards and Technology, USA.
- Rok Bojanc & Borka Jerman-Blazic (2008). „An economic modeling approach to information security risk management", *International Journal of Information Management 28*.
- Sonnenreich, W. (2006), „Return On Security Investment (ROSI): A Practical Quantitative Model", *Journal of Research and Practice in Information Technology, Vol. 38,* No. 1, a publication of the Australian Computer Society, Australia
- Schneier, B. (2004). *Secrets & lies: Digital security in a networked world*. Indianapolis, Wiley Publishing.
- Whitman, M. E. (2003), „Enemy at the gate: Threats to information security", *Communications of the ACM, 46(8).*

APPLYING STRING EDIT DISTANCE TO DISCOVER BUSINESS SUBSIDIARIES

HANCU Lucian
Ph.D., Faculty of Economic Sciences and Business Administration/Department of Business Informatics, "Babeş-Bolyai" University, Cluj-Napoca, Romania, Email: lhancu@softproeuro.ro

Abstract: *Mergers and Acquisitions models for predicting future strategic transactions between two business entities require high-quality Information Systems, with precise data from multiple sources. Discovering whether a company is a subsidiary of a stronger one is a crucial task for correctly predicting future transactions among the available companies. In this article, we apply string edit distances in order to discover the main company of a business subsidiary. These findings have multiple applications, ranging from improving the strategic models for predicting Mergers and Acquisitions transactions, to correctly input the business partners' financial information on each of the business entities Information Systems.*

Key words: Mergers and Acquisitions; Information Systems; Business Organization; String Edit Distance

JEL classification: G34, O12, C44

1. Introduction

Modelling mergers and acquisitions transactions is a challenging task when using data extracted from multiple web-based Information Systems. In this approach, the quality of the data extracted from each one of the information system is crucial, applying more information into the model, if inappropriately used, having a bad impact on the final results (Hancu, 2012).

In this paper, we focus on extracting business subsidiaries relationships of the type subsidiary and main company from a set of subsidiaries candidates. The approach comes in the context of a need for more advanced techniques that would be used in order to add semantic meaning to business transactions in particular and business-related data – in general. These transactions are, in our paper's context, the business transactions between suppliers and clients that have to be reported periodically by the business entities among the Romanian economy.

The entering of a business transaction in the case of a small entity made with a small entity is extremely simple. The partners know each of the other party well, as in most of the cases are long=term traditional partners. What if the transaction is made between a large and a small company? Or between two large companies? The transaction is made between the subsidiaries of the same city, for instance, but the invoices are released by the main companies' financial departments, which might be in the same city, or anywhere in the country. In this context, correctly introducing the business transaction into the Information Systems of both companies can be a challenge. One company, for instance, could input the subsidiary as the partner, or even both companies can input the subsidiary into the system.

A still more challenging situation is when the invoice of the partner contains two identifications of companies. One would be the main company's details; the other – the subsidiary information. Our experience in verifying and correcting VAT-related declarations has shown that most of the people input the subsidiary as the input company for the transaction, as they in fact do the transaction with the subsidiary, but are not aware of the fact that the main company resides in a totally-different city with totally different financial information and attributes (the main problem usually relies in the fact that the main company is VAT payer, while the subsidiary is not the VAT payer).

In this context, applying automated techniques in order to extract relationships of the form business subsidiary - main company is a challenge, more companies not having the sufficient time nor the sufficient interest in verifying their portfolio of business partners, which in fact can yield to unexpected surprises. This article describes how to apply string edit distances techniques in order to discover which one of the potential candidates are the main companies of the business subsidiaries identified as a list of business partners.

When dealing with solving all the business relationships of the type business subsidiary, having the whole information on the active companies is among the highly-ranked success factors. In this context, we use only a small portion of business partners that were priory classified as non-VAT payers –

thus potential business subsidiaries and a larger database of active companies which are VAT payers. The problem shall be stated in the subsequent section of the paper.

The paper is organized as follows. In the next section of the paper, we thoroughly describe the problem statement: discovering business subsidiaries among a list of business entities. In the third section, we point out the algorithm of edit distance that we apply from the research literature, then present the results of our experiments. In the fifth section, we point out a series of issues regarding the performed experiments and sketch some possible future developments. In the last section, we conclude by reviewing the topics discussed in this paper and set directions for future work.

2. Problem Statement

Discovering business subsidiaries among a list of business partners is a very common nowadays business problem among the Romanian (and not only) active business entities. The problem is defined as follows: Having a set ***S*** of business entities which are not VAT payers, discover a set of pairs (***s***, ***m***), where ***s*** is a member of ***S*** and ***m*** is the member of the VAT payers set, and, in addition, ***s*** is a subsidiary of ***m***.

Usually, this problem is not known by the business entities having it. They assume that the ***s*** is the VAT payer and thus report through their financial reports this inexact information. The first condition that must be true in order to correctly diagnostic the problem is that the business entities would be willing to periodically check the portfolio business partners against various public Information Systems so as to derive the list of the VAT payers and the list of non-VAT payers. The second list would be subject of interest to our research.

Generally speaking, when we have a set of business entities pairs (s, m) of the form expressed above, not always ***s*** is a non-VAT payer and ***m*** is a VAT payer. There are cases in which both ***s*** and ***m*** are non-VAT payers, as in the case of non-profit organization or the case of both ***s*** and ***m*** are VAT payers (for instance, some financial institutions).

There are several motivations that convinced us in finding and applying automatic algorithms for solving the above problem. A subset of them is depicted and discussed below:

- **Space of the problem**: when the list of business partners is high above a specific threshold, the business entities do not dedicate time and effort to checking whether their business partners are VAT payers or not-VAT payers. This non-verification of business partners comes also from the culture of the organization: there are organizations in which their internal procedures state that prior to enter one partner into the database several checks are performed in order to derive whether it is a VAT partner or not, whether the company is valid or has several financial problems.
- **Overconfidence**: the business partners are overconfident in their Information Systems, lacking the fact that the environment is extremely dynamic: a company that was VAT payer one year ago can be suspended, thus lacking the VAT attribute.
- **Lack of sufficient time**: lots of times, when computing the financial declarations to the public financial institutions, the business partners use the last days before the deadline in completing these declarations, thus altering the quality of the entered information.
- **Unavailability of public Business Information Systems**: from our experience of several years in monitor and correct the partners' portfolio of clients, the periods just before the deadlines of the fiscal declarations were also the periods in which the public business Information Systems were unavailable, due to overcharge them with both manual and automatic queries.
- **Robot exclusions of public Business Information Systems**: related to the previous argument, nowadays Information Systems use robot-exclusion software that drastically limit the access of computer software to the information of the financial entities, thus limiting the access of automatic agents to these Information Systems and deteriorating the quality of the information regarding the business partners of the business entities.

Analyzing several hundreds portfolios of clients from more than 5 years of financial declarations (from 2007 up to date) we have empirically discovered that a large amount of non-VAT payers are subsidiaries of other VAT payers. Our aim was, since then, to derive algorithms in order to automatically point out the pairs (***subsidiary***, ***mother_entitity***) and perform automatic corrections to the portfolios of business partners.

From the economic crisis point of view, the problem is of great importance. Two potential implications contribute to the motivations of applying the above-stated research. First, periodically check

the portfolios of business partners can save the clients from bad surprises – partners that are not more active, partners that have filled for insolvency, or still worse – for bankruptcy.

The other implication – more positive – would be to correct the portfolios of clients in order to further perform cross verifications between the transactions reported by each one of the business partners. In the following sections we shall also discuss the issues and the advantages of the proposed techniques, in the economic context – in general and in the context of the prolonged economic crisis – in particular – in which re-establishing the trust among the business partners is crucial, so is the re-establishing of the reputation for an inactive business partner which re-entered business life.

In the following section, we depict the notions of edit distance and describe the algorithm that we use in order to correctly classify the business entities into subsidiaries of other business entities and point out which is the correct name of the ***mother entity m***.

3. Applying String Edit Distance Techniques

Giving a set of non-VAT payers ***S***, and a set of VAT payers ***V***, the aim of this research is to apply automatic algorithms in order to derive the pairs (***s***, ***m***), where ***s*** is a member of ***S***, ***m*** a member of ***V*** and ***s*** is the subsidiary of ***m***. As we shall see in the following section, the subset of ***s*** members of the pairs (***s***, ***m***) is only a small subset of ***S***, there are a lot of companies that are not-VAT payers and are not a subsidiary of any other business entity.

With this purpose in mind, we apply String Edit Distance Algorithms in order to minimize the edit distance and to predict which candidate has the chance of being the mother company of the subsidiary candidate *s*. If such a company is revealed, it is pointed out the pair (***s***, ***m***).

Levenshtein (1965), cited in Cenk Sahinalp (2005), describes a dynamic programming algorithm for computing the string edit distance between two input strings in polynomial time O(n * p), where n is the length of the first string and p is the length of the second string. Even if other approaches optimize this algorithm by applying various methods, the most scalable algorithm is not the subject of our research. We recall that the edit distance between two strings ***s*** and m is the minimum number of edits – being either insertions, deletions or replacements – that are necessary to transform the ***s*** in ***m***.

In this paper, we plan to use the Levenshtein String Edit Distance in order to discover the subsidiaries of companies and output a subset (s, m), where s is the subsidiary and m is the mother business entity. We describe below the main steps of our algorithm (see Table 1):

Table 1: Algorithm for Minimizing the Edit Distance btw. Subsidiaries and Main Entities

Step	Description
1.	Derive the list of the non VAT payers ***S***
2.	For each ***s*** in ***S*** that contains **S**ubsidiary Substring **ID**s (**SSID**)
3.	Compute the subset of subsidiary candidates of Cleansed String
4.	Verify whether one of the candidates up to level k is the correct one
5.	Output the number of true positives (TP)

The algorithm begins with the computation of the set of non-VAT payers ***S*** by use of automatic or semi-automatic tools. A thorough discussion on the topics used in this step of the algorithm can be found in (Hancu, 2007) and is far beyond the scopes of this article. Upon the identification of the set ***S***, we take every member of the set that contains the subsidiary substring IDs (in Romanian language we use FILIALA, DIRECTIA and SUCURSALA as the most common used substring identification marks for a subsidiary company, both public and private).

Later on, during the third step, we delete the identification subsidiary substring and the substring that follows it, and then compute a set of company candidates by applying the string edit distances between the cleansed string and companies from the VAT-payers database. The list of candidates from the VAT-payers database is later sorted according to the edit distance between each one of the candidates and the cleansed string.

The fourth step of the algorithm applies a cross-verification technique. Prior to using this algorithm for the experiments, we have manually computed a list of pairs (subsidiary, mother_entity) by using the same edit distance techniques, but leaving the final decision (whether two companies are in a subsidiary relationship) to humans. Now, we take the list of ordered candidates and verify up to level k whether the mother entity from the manual classification lies between them.

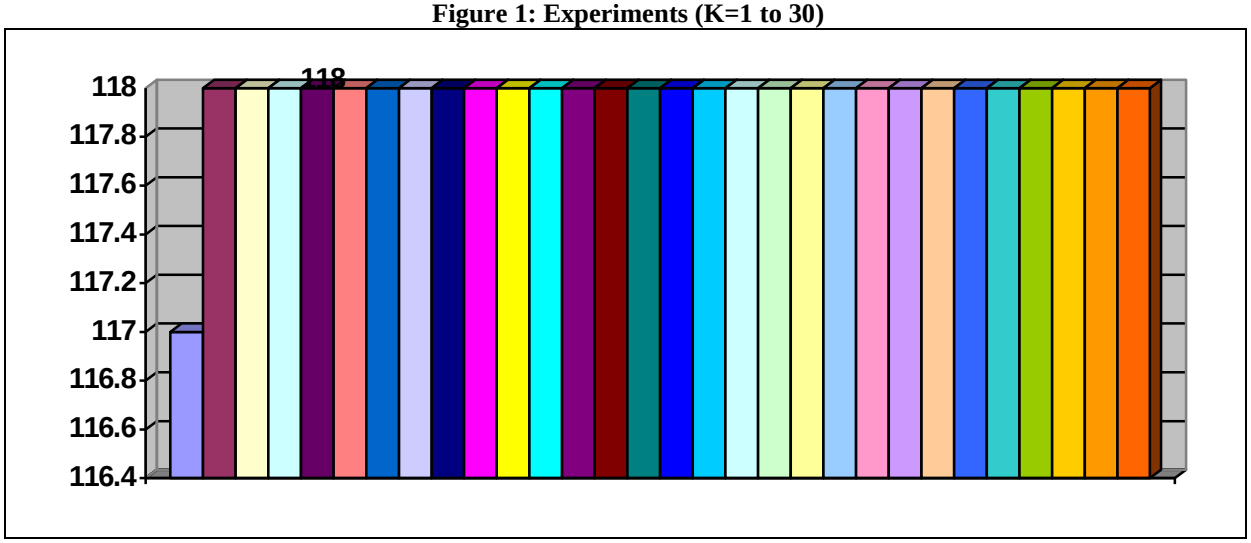
The algorithm continues with the output of the true positives, i.e. the number of companies correctly classified as mother companies of a subsidiary, where the candidates were searched up to the level k of the sorting by using the edit distance between the candidate and the cleansed string.

In the next section we depict the set of experiments made with the data, whereas the forthcoming section describes several issues that we discovered during our experiments. We also derive and explain several advantages of applying automatic and semi-automatic cleansing of subsidiaries.

4. Results

In the **Figure 1** we depict the results of our experiments and here we point out our conclusions. We have applied string edit distance algorithm to a set of non-VAT payers containing 23671 entities, from which we have identified a list of 290 candidates which contained the substring subsidiary IDs that we apply in the algorithm. From this set, 240 pairs of ***(subsidiary, mother_company)*** have been manually identified.

Figure 1: Experiments (K=1 to 30)



The experiments were performed as follows: for each experiment we considered a threshold (ranging from 1 up to 30) in which the companies' candidates were analyzed in order to determine whether one of them can be classified as the mother company of the input cleansed string. The experiments from the level 2 up to the 30 level gives us a surprising stable results, which encourages us in further applying these techniques. The stability of the algorithm demonstrates that *if a subsidiary is found, the edit distance between the cleansed string and the name of the company is among the two lowest edit distances between the candidate strings*.

By manually inspecting the initial database of non VAT payers we have identified a large number of business entities which were state-owned institutions, which were not considered during our experiments. While companies' subsidiaries have a more suggestive name including the main company name and the location of the subsidiary together with one of the subsidiaries substrings that we use in our algorithm, the state-owned institutions can have a totally different name when analyzing the local institution and the central institution, thus applying several edit distance measure would be inappropriate. For these state-owned institutions, more advanced techniques would need to be developed in order to discover subsidiaries in an automatic or semi-automatic manner.

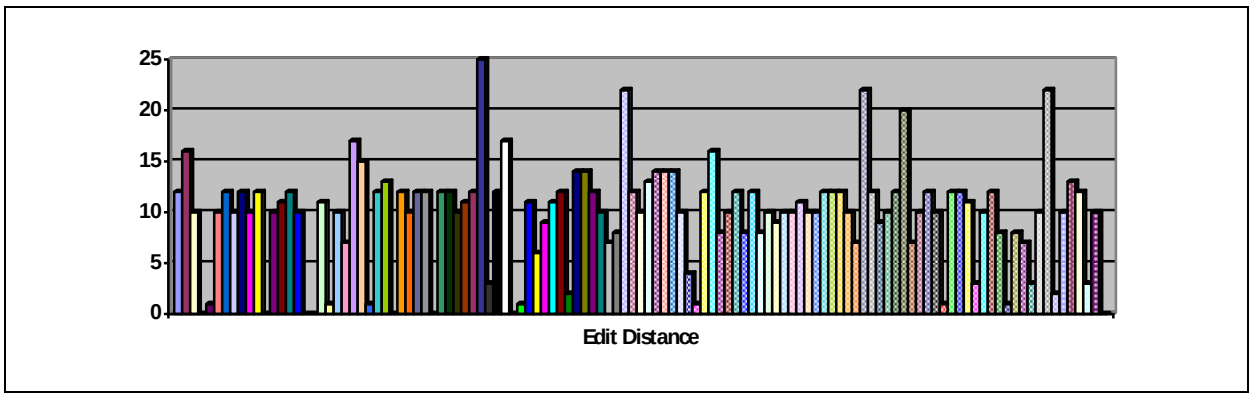
On the other part, by inspecting the manual database of computed pairs of subsidiaries and mother companies, some interesting facts were discovered, which demonstrate the potential of the techniques presented here. Among the list of mother companies, several business entities were state-owned agencies (which were not-VAT payers), while others were private agencies or non-profit organizations, also non-VAT payers. As we have stated above, the database we use for checking the companies name contain only VAT payers companies, thus not finding non-VAT mother entities is natural.

When analyzing the total number of potential subsidiaries and the number of pairs found (240) and the relative number of subsidiaries found in the analyzed database (above 1,04%) one would say that

the percent is too small. The fact that a database contains only a small number of subsidiaries to be identified does not mean that we would have to leave them erroneously registered into the database.

From the other point of view, considering that we do not apply these techniques in discovering companies subsidiaries: having 240 errors into our database means that the quality of the information can be seriously affected. In contrast, our idea was to use and develop semi-automatic techniques in order to correctly find business main companies from the business subsidiaries' names.

Figure 2: Edit Distance between Subsidiary and Main Company: Experiment (K=1)



In the **Figure 2** we analyze the edit distances between the subsidiaries found during our experiment with K=1 (the number of true positives being 117, the K=2 and the other experiments producing only a slightly improved result). We observe that the average edit distance between the subsidiary and the main company string name is between 7 and 10 characters, with few differences greater than 20 and several cases in which the edit distance is less than 5.

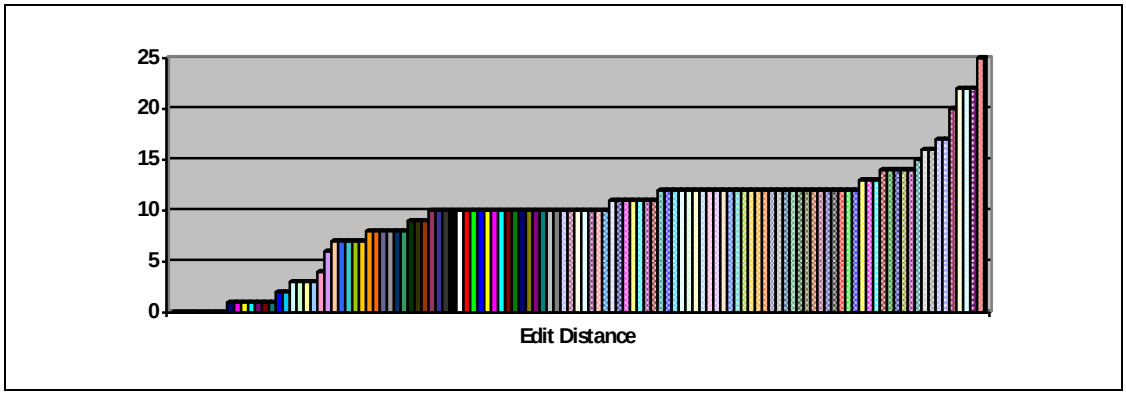
This later finding can be easily observed from the **Figure 3**, in which we depict the same data as in the **Figure 2**, but in a sorted order by the edit distance between the candidate string and the main company string. We explain this finding as the subsidiary contains the name of the main company, the city of residence of the main company (the most used being BUCURESTI) and then the subsidiary identification followed by the city, county or region in which the subsidiary is located. Considering that the substring containing the subsidiary ID and the substring following that ID are truncated prior to applying the edit distance measures, we conclude that the edit distance contains the city of residence of the main company.

A future improvement of these techniques would be to apply data cleansing of the main cities that are included into the candidate string in order to improve the results and diminish the edit distance between the candidate string and the main company's name. A difficulty to this task would be to know whether the main city name is included or is not included in the original name of the company.

While the most found cases of main companies' names do not include their city of residence into their name, we cannot consider that a general rule, as there are several companies which include the name of the city of residence into their name, this being in fact the choice of the company's founders and, thus, a variable with a high number of instances.

During the next section, we depict several findings of our research – the limits (or issues) together with the advantages (the tough points) of applying the automatic techniques in discovering business subsidiaries. We also discuss the implications of the current economic crisis on the quality of the input data into the Information Systems of the business entities and the importance of the periodically checking on the current status of the business entities, as market conditions change rapidly, and so the state of the business partners – from active entities to inactive – and vice-versa.

Figure 3: Edit Distance between Subsidiary and Main Company: Experiment (K=1) (sorted)



5. Discussions

When analyzing the results of our experiments, a series of issues were discovered, that we analyze below. These issues negatively affected the results of our experiments and are still under investigations for future developments of our researches.

- **Both the subsidiary and the main company are not VAT payers**. Usually restricted to a small number of individuals, the cases when both the subsidiary and the main company are both non-VAT payers affect the results of our experiments. Since the database in which we search for main companies contain only VAT payers, this latter case would be impossible to be solved automatically, a database containing non-VAT payers being required in such case. These cases apply for instance when the subsidiary and the main company are non-profit organizations, which do not pay VAT, or are state-owned agencies, also non-VAT payers.
- **The Company A is a subsidiary of the company B**, but the company **B** has *changed its name*, after a business transaction of the type merger and acquisition. This comes in correlation with the fact that the company A is not more active to the market, while the company B is the one that is now active, after the transaction took place (usually the main company of A and the B company have merged, or B acquired the main company of A). In this case, applying string edit distance algorithms to the revealing of the main company having the subsidiary A company is difficult; the edit distance between the A company and the B company being in most of the cases higher than the edit distances between the A company and other available business companies, which are VAT payers, but have no link to the initial A company
- **The name of the main company does not contain the name of the town of residence** of the company, but the name of the company does contain the town name. This also limits the results of the technique for deriving whether two companies are of the form subsidiary-main company. An idea to overpass this situation would be to clean the names of the towns from the names of the companies and apply the edit distances after such cleansing phase. The problem still remains, as we cannot know if the name of the main company has or has not the name of its town (or county of residence) contained in its name.

Even if these issues affect the quality of the results of applying automatic techniques in order to predict whether one company is the subsidiary of another company, by applying edit distances measures to both companies' names, a series of *advantages* contribute to the importance of our research:

- **Improving Mergers and Acquisition models for predicting strategic transactions based on the financial statements analyzes of the business companies:** when analyzing the business indicators of the subsidiary company and the main company, we might discover that the main company's financial indicators are better than the subsidiary's financial indicators: the main company has a higher turnover than the subsidiary, its tangible and intangible assets are higher, the market power of the main company is better compared with the subsidiary's market power, so is the number of the employees. This might conduct, normally, to the reveal of an acquisition recommendation from the strategic model, which would not be correct, as the small company is already a subsidiary of the main company.
  The model of mergers and acquisitions strategic decisions should also take into consideration that there is a subsidiary relationships between the two companies, thus output a different decision score, for instance, *Conservation* – or no strategic action, for the decision model, according to Hancu (2008).

A thorough discussion on the topics of Mergers and Acquisitions can be found in (DePamphilis, 2011). One of the main idea that was employed from the cited research book is that a thorough analysis and valuation must be performed prior to deciding upon a Merger or Acquisition transaction – thus establishing the subsidiary relationships comes in this context.

- **Improving the Information Systems of the Business partners:** from the business entity's perspective, having a list of several tens of thousands of business partners that have to be checked continuously in order to determine whether they are still valid entities, whether they have been bought by other entities, is a difficult task. Thus, deriving semi-automatic techniques for keeping such database systems up-to-date is crucial. In the case that the public Information Systems apply human verification mechanisms for disallowing computer robots to access such information, the problem of keeping the business partners' database up-to-date is still more difficult. Thus, applying automatic and semi-automatic techniques that monitor the Information Systems remains challenging, more automatic and semiautomatic tool still being needed for the correct extraction of data form such Information Systems.
- **Improving the centralized cross-verification Information Systems:** when applying cross-verification techniques in a centralized Information System, the quality of the data is crucial. When one party uses the mother (main) company for the reporting of transactions that are under the subject of cross-verifications, and the other uses the subsidiary entity, abnormalities appear in the cross-verification database. Thus, a semiautomatic or an automatic cleansing would be needed, prior to perform cross-verifications.

A novel advantage in applying the techniques described above is the discovery of mergers and acquisitions transactions that were not known prior to use these techniques. As we have stated above, one of the limits of the automatic algorithm is that it does not deal with Mergers and Acquisitions transactions, where a change of a name of company appeared. On one part, this is a disadvantage of the proposed method. On the other part, the instances that were not correctly classified by the system and had to be manually analyzed had a major number of mergers and acquisitions transactions included into the life of the involved entities. This means that after an M&A transaction the company has changed its name according to the new owners' business trademarks.

When considering the implications of these database errors on the global and national crisis, we state that by applying error-free cross-verifications on the transactions between the two companies (the transactions of A to the company B reported by A being equal to the transactions of A to the company B reported by B would have to be equal) the trust among the business participants increases. When a company is found to have all the cross-verifications correct, its reputation among the participants to the market would increase. This statement comes in the context of the crisis, when several business partners find appropriate to use less-correct behaviour in their financial presence to the market, behaviour which in the case of cross-verifications would be easily to be identified and presented to other companies of the market.

From the other point of view, in extremely tense periods, people at work can feel high level of stress, which can affect their quality of work, according to their personality traits (Goleman, 1999), (Littauer, 2011). While some of them would benefit from short-term high-stress levels, most of the others will feel that the quality of work (and life) decreasing. In this context, the quality of the work can be corrected totally or partially by automatic or semiautomatic techniques, some of them being described in this paper.

The benefits of having a correct Information System can easily surpass the cost of having such high-quality database, both in term of time and financial resources. Not to mention that having a high-quality Information System would contribute to obtaining more accurate Business Intelligence reports (Michalewicz, 2006), which would be of great benefit of both organizations and their stakeholders.

6. Conclusions

In this paper, we have presented a technique for applying string edit distances to a database of available companies that were previously classified as non-VAT payers (thus, possible subsidiaries) in order to automatically discover whether they are subsidiaries or not of one of the available VAT payers companies. The result is promising, as the problem is common for the companies having several tens of thousands of business partners and that do not regularly check on whether their business partners are still valid payers or not. For such companies, applying semi-automatic techniques remains the sole solution for keep their database as clean as possible.

In the context of the new informative declarations that have to be submitted periodically to the financial national authorities, having a clean (error-free) database is a normal desiderate, applying such automatic techniques together with a human cross-validation of the results assures a better quality of the data contained into the Information Systems of the companies.

This research can be further developed in several directions. In this paper, we have applied the Levenstein Edit Distance between two strings. A further development would be to apply other edit distance algorithms and compare the results. Other further development would be to apply more cleansing techniques in order to pre-process the name of the subsidiary, for obtaining a better input for the edit distance score. One suggestion would be to apply cleansing techniques for the town names that are contained within the company's name, but we expect to encounter exceptions that these town (or geographical names) are really part of the main company's official name.

Other further improvements would be to consider other substring subsidiary IDs. Increasing the number of used IDs would be a hard task, since a larger database of companies would be required. The fact that more IDs are introduced into the algorithm increases the chances that the algorithm considers a candidate for the edit distance measures, and thus it increases the chance to find a candidate for the main company by minimizing the edit distance between the two strings.

All these further developments would improve the quality of the expected results, thus demonstrating the high potential of the discussed techniques in automatically identifying business subsidiaries from a wide portfolio of business partners.

7. Acknowledgement

This work has been partially financed by the CNCS-UEFISCDI under the grant PN-II-RU PD397 / 2010 – Collaborative Methods for Improving Mergers Decisions' Support. This research is the result of more than five years of manual and semi-automatic corrections performed on the portfolios of clients. We thank the numerous business partners all around the country for giving us the possibility to derive and implement more advanced algorithms for solving their problems.

8. References

- Cenk Sahinalp, S. (2008) *Edit Distance under Block Operations*, in *Encyclopedia of Algorithms*, edited by Ming Yang Kao, Springer-Verlag, 2008, pp.255-257.
- DePamphilis, D., M., (2011) Mergers, Acquisitions and other Restructuring Activities, Academic Press, Elsevier, 2012.
- Goleman, D. (1999), *Working with Emotional Intelligence*, Bloomsbury Publishing, 1999.
- Hancu, L. (2007) *Enhancing the Invisible Web*, in Proceedings of the International Conference KEPT 2007, Cluj-Napoca, Romania, 2007.
- Hancu, L. (2008) *Pruning Decision Trees For Easing Complex Strategic Decisions*, in Proceedings of the 2008 Business Information Systems Day, Tiberiu Popoviciu Seminars, Vol. V(2008), MediaMira Printing House.
- Hancu, L. (2012), *Mining Strategic Decisions Information Systems for Predicting Future Market Concentrations*, in Proceedings of the International IADIS Information Systems Conference, Berlin, Germany, March 2012.
- Levenshtein, V. I. (1965) Binary *codes capable of correcting deletions, insertions, and reversals*, Doklady Academii Nauk SSSR, 163(4), pp. 645-848.
- Littauer, F. (2011), *Personality Plus at work, How to work successfully with everyone*, Revell, 2011.
- Michalewicz, Z., Schmidt, M., Michalewicz, M., Chiriac C., Adaptive Business Intelligence, Springer Verlag, 2006.

SIMULATING THE NUMBER OF ROMANIAN'S PUPILS AND STUDENT USING ARTIFICIAL INTELLIGENCE TECHNIQUES AND COUNTRY'S ECONOMIC ESTATE

ILIE Constantin
PhD. Eng./Lecturer, Faculty of Mechanical, Industrial and Maritime Engineering, OVIDIUS University from Constanta, Constanta, Romania, ilie.ctin@yahoo.com

ILIE Margareta
PhD. Ec./Lecturer, Faculty of Economic Science, OVIDIUS University from Constanta, Constanta, Romania, udrescu_margareta@yahoo.com

MELNIC Lucia
PhD. Eng./Professor, Faculty of Mechanical, Industrial and Maritime Engineering, OVIDIUS University from Constanta, Constanta, Romania, lucia_melnic@yahoo.com

Abstract: *The authors present the result of a research that uses artificial neural networks in order to simulate the number of Romanian's pupils and students (PAS), considering the country's economic situation. The objective is to determine a better method to forecast the Romanian's PAS considering its nonlinear behaviour. Also the ANN simulation offers an image about how inputs influence the PAS. In conclusion, the use of the ANN is considered a success and the authors determine the possibility that ANN research application be extended to other countries or even to the European zone.*

Key words: Artificial intelligence, Artificial Neural Network (ANN), Pupils and Students, GDP per capita

JEL classification: C45

1. Introduction

The importance of the government implication in the education is more emphasised with the spectre of financial crisis effects. But the government's investment in education, research and jobs is determined by its economic result, such as Gross domestic product (GDP). Even if it is hard to find a linear or direct link and determination between the GDP and the number of the Romanian's pupils and student (PAS) a country has, it is obvious the existence of a direct relation between them. The present paper offer a method that uses nonlinear mathematics to determine the influence between indices for economic activity and PAS in order to assure the base for future research and to forecast the PAS accordingly to country's GDP. The mathematical method consists in Artificial Neural Networks (ANN) a branch of Artificial Intelligence, presented in subsection 2.1.

The researchers proposed that ANN should learn from the past PAS values and forecast accurate values of the PAS determining also nonlinear connections between the PAS and the economic elements defined in subsection 2.2.

Objectives of the research are:

1. Building, training and validating a specific ANN for the simulating of the PAS considering the measures for economic activity and the imposed conditions;
2. Testing the trained ANN in order to check that the difference between the real data and the simulated is smaller than 5%;

2. Theoretical elements

2.1. Artificial neural network (ANN)

Today, Artificial Neural Network (ANN) can be found in the majority of the human life activity, their potential being considered immense. Starting with the human brain and continuing with the actual common application of the ANN, their use effectiveness is demonstrated as can be seen in Akira Hirose's book "Complex-valued neural networks: theories and applications" (Akira, 2003).

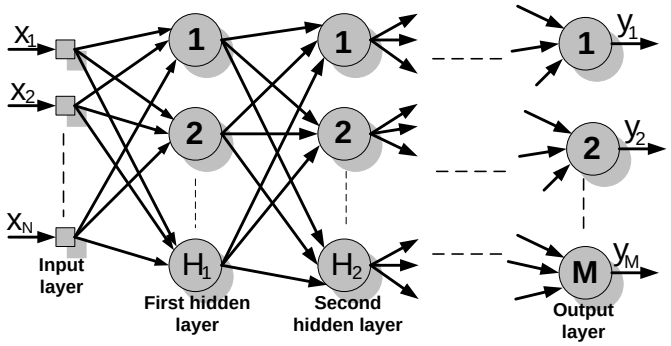
Regardless of the type, ANN has a few certain common unchallengeable elements mentioned by Zenon WASZCZYSZYN in his book: "Fundamentals of Artificial Neuronal Networks" (WASZCZYSZYN, 2000):

- micro-structural components: processing elements – neurons or nodes

- input connections of the processing elements
- output connections of the processing elements
- the processing elements can have optional local memory
- transfer (activation) function, which characterizes the processing elements.

The schematic graphic representation of a multilayer ANN is presented in figure 1, were x_i ($i=\overline{1,N}$) is the input ANN values, H1 and H2 are the numbers of neurons from the first and second hidden layer, M is the number of neurons from the output ANN layer and y_j ($j=\overline{1,M}$) is the values of the output data.

Figure 1: Schematic representation of a multilayer ANN.



Source: (WASZCZYSZYN, 2000)

2.2. Input and Output data

The input and output data can be found on European Commission website (http://epp.eurostat.ec.europa.eu) and they are are defined as follows:

GDP per capita in PPS (GDP), as - Gross domestic product (GDP) is a measure for the economic activity. It is defined as the value of all goods and services produced less the value of any goods or services used in their creation. The volume index of GDP per capita in Purchasing Power Standards (PPS) is expressed in relation to the European Union (EU-27) average set to equal 100. If the index of a country is higher than 100, this country's level of GDP per head is higher than the EU average and vice versa. Basic figures are expressed in PPS, i.e. a common currency that eliminates the differences in price levels between countries allowing meaningful volume comparisons of GDP between countries.

Share of government budget appropriations or outlays on research and development (GOVB), as % of total general government expenditure - Data on Government Budget Appropriations or Outlays on Research and Development (GBAORD) refer to budget provisions, not to actual expenditure, i.e. GBAORD measures government support for R&D using data collected from budgets. The GBAORD indicator should be seen as a complement to indicators based on surveys of R&D performers, which are considered to be a more accurate but less timely way of measuring R&D activities. In this table, total GBAORD is expressed as a percentage of total general government expenditure.

Real GDP per capita (RGDP), growth rate and totals as percentage change on previous year - Euro per inhabitant - GDP includes goods and services that have markets (or which could have markets) and products which are produced by general government and non-profit institutions. For measuring the growth rate of real GDP, the GDP at current prices are valued in prices of the previous year and the thus computed volume changes are imposed on the level of a reference year; this is called a chain-linked series. Accordingly, price movements will not inflate the growth rate. Real GDP per capita is calculated as the ratio of real GDP to the average population of a specific year. It is often used as an indicator of how well off a country is, since it is a measure of average real income in that country. However, it is not a complete measure of economic welfare. For example, GDP does not include most unpaid household work. Neither does GDP take account of negative effects of economic activity, like environmental degradation.

Employment rate by sex (EMPR), as % - This employment rate is calculated by dividing the number of persons aged 15 to 64 in employment by the total population of the same age group. This is not the Europe 2020 employment rate indicator which refers to persons aged 20 to 64. The indicator is based

on the EU Labour Force Survey. The survey covers the entire population living in private households and excludes those in collective households such as boarding houses, halls of residence and hospitals. Employed population consists of those persons who during the reference week did any work for pay or profit for at least one hour, or were not working but had jobs from which they were temporarily absent.

Pupils and students (PAS), in Thousands - This table includes the total number of persons who are enrolled in the regular education system in each country. It covers all levels of education from primary education to postgraduate studies (excluding pre-primary education). It corresponds to the target population for education policy.

3. Simulating the PAS using ANN

Regardless of the ANN used, the processed data and the simulated problem the phases in the implementation of ANN are the same (Ilie, 2010, p. 667), as follows.

3.1. Phase 1 – Initial data analysis.

Data analysis consists in dividing data into separate columns, defining types of these columns, filling out missing number values, defining the number of categories for categorical columns etc.

Data analysis revealed the following results:

- 5 columns and 10 rows were analysed;
- Data partition method: random;
- Data partition results: 8 records to Training set (80%); 1 records to Validation set (10%) and 1 records to Test set (10%).

As can be seen, the database is divided in three sets. While the training sets are used only for training, the validation and testing sets are used also for testing. Presented in table 1 are the maximal and minimal value limits of the input and output data.

Table 1: The Database

Data	Parameter name	Format	Max	Min
Input data	GDP per capita in PPS (GDP)	Numeric	42	26
	Share of government budget appropriations or outlays on research and development (GOVB), %	Numeric	0.36	1.01
	Employment rate by sex (EMPR), %	Numeric	57.6	64.2
	Real GDP per capita (RGDP), %	Numeric	-1.9	8.8
Output data	Pupils and students (PAS), x1000	Numeric	3831	4019

Phase 2 – Data pre-processing.

This is the phase in which the above defined data are prepared and reshaped for an easier use and for obtaining best results, according to the requirements and the imposed results.

Pre-processing means the modification of the data before it is fed to a neural network. Pre-processing transforms the data to make it suitable for neural network (for example, scaling and encoding categories into numeric values, “one-of-n” or binary) and improves the data quality (for example, filtering outliers and approximating missing values), as different software uses different methods (www.alyuda.com).

Numeric encoding means that a column with N distinct categories (values) is encoded into one numeric column, with one integer value assigned for each category. For example, for the Capacity column with values "Low", "Medium" and "High", "Low" will be represented as {1}, Medium as {2} and High as {3} (www.alyuda.com). In table 2, the characteristics of pre-processed data are shown.

Table 2: The Pre-processed database.

Parameter	Scaling range	Min	Max	Mean	Standard deviation	Scaling factor
GDP	[-1..1]	-1	1	31.6	5.2	0.12500
GOVB		-1	1	0.612	0.25	3.07692
EMPR		-1	1	60.09	2.608	0.30230
RGDP		-1	1	4.67	3.540	0.18692

PAS	[0..1]	-1	1	3921.42	63.88	0.00529

The results of completed pre-processing process are:

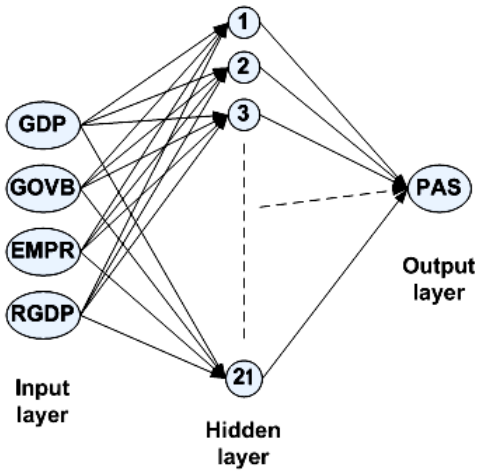
- Columns before preprocessing: 5;
- Columns after preprocessing: 5;
- Input columns scaling range: [-1..1];
- Output column(s) scaling range: [0..1];

Phase 3 – Artificial neural network structure.

Considering the characteristics of the simulated process, many ANN structure can be determined and compared through the specificity of the process and the data that are being used for simulations. Thus, the feed forward artificial neural network is considered the best choice for present simulation (Barbeş 2009, p. 301).

After building and testing several ANN with feed forward structures, having in mind the comparison of the errors between the real data and ANN output data, the best ANN network was defined. This has the following structure (figure 2): 4 neurons in the input layer, 21 neurons in the hidden layer and 1 neuron in the output layer.

Figure 2. Simplified graphic representation of ANN with structure 4-21-1.



Also, the activation function for different layer neurons are chosen: Hidden layers activation function is hyperbolic tangent - ratio between the hyperbolic sine and the cosine functions (or expanded, as the ratio of the half-difference and half-sum of two exponential functions in the points z and –z) - and Output layer activation function is logistic (sigmoid curve) (Năstac, 2000).

Phase 4 – Training.

Being an essential phase in the use of ANN, the training must use certain training algorithms which essentially modifies the structural elements of ANN (weights) modified through several iterations. Those modifications establish the future ANN accuracy. For the selected ANN, the most common training algorithm is Back propagation algorithm.

Back propagation algorithm is the best-known training algorithm for multi-layer neural networks. It defines rules of propagating the network error back from network output to network input units and adjusting network weights along with this back propagation. It requires lower memory resources than most learning algorithms and usually gets an acceptable result, although it can be too slow to reach the error minimum and sometimes does not find the best solution (www.alyuda.com). For a quicker training process, a modification of the back propagation algorithm, called quick propagation, was made.

Quick propagation is a heuristic modification of the back propagation algorithm invented by Scott Fahlman. This training algorithm treats the weights as if they were quasi-independent and attempts to use a simple quadratic model to approximate the error surface. In spite of the fact that the algorithm has no theoretical foundation, it is proved to be much faster than standard backpropagation for many problems. Sometimes the quick propagation algorithm may be instable and inclined to block in local minima (www.alyuda.com).

The training conditions were established in order to achieve the best results using the quick propagation training algorithm. Thus, the quick propagation coefficient (used to control magnitude of weights increase) was 0.9 and the learning rate (affects the changing of weights – bigger learning rates cause bigger weight changes during each iteration) was 0.9. The small values of these two, especially of the learning rate, are explained by the necessity of avoiding the local minima blockage.

The results of training details are presented in table 2.

Table 2: The Training details.

Retrain	Iters.	Train error	Validation error	Test error	AIC	Correlation	R-Squared	Stop reason
1	$1.1x10^6$	0.01	0.421462	8.699985	200.5231	1	1	Desired error achieved
Summary								
Min	n/a	0.01	0.421462	8.699985	200.5231	1	0.999029	n/a
Max	n/a	0.01	0.421462	8.699985	200.5231	1	0.999029	n/a
Average	n/a	0.01	0.421462	8.699985	200.5231	1	0.999029	n/a
Std. dev	n/a	0	0	0	0	0	0	n/a

- AIC is Akaike Information criterion (AIC) is used to compare different networks with different weights (hidden units). With AIC used as fitness criteria during architecture search, simple models are preferred to complex networks if the increased cost of the additional weights (hidden units) in the complex networks do not decrease the network error. Determine the optimal number of weights in neural network;
- Iters. are the iterations;
- R-squared is the statistical ratio that compares model forecasting accuracy with accuracy of the simplest model that just use mean of all target values as the forecast for all records. The closer this ratio to 1 the better the model is. Small positive values near zero indicate poor model. Negative values indicate models that are worse than the simple mean-based model. Do not confuse R-squared with r-squared that is only a squared correlation;

The stop training conditions were: maximum of 1500000 iterations or a maximum absolute training error value of 0.001. The results of training are Number of iterations: 1105307 (Time passed: 00:02:38 min.).

3.2. Training results

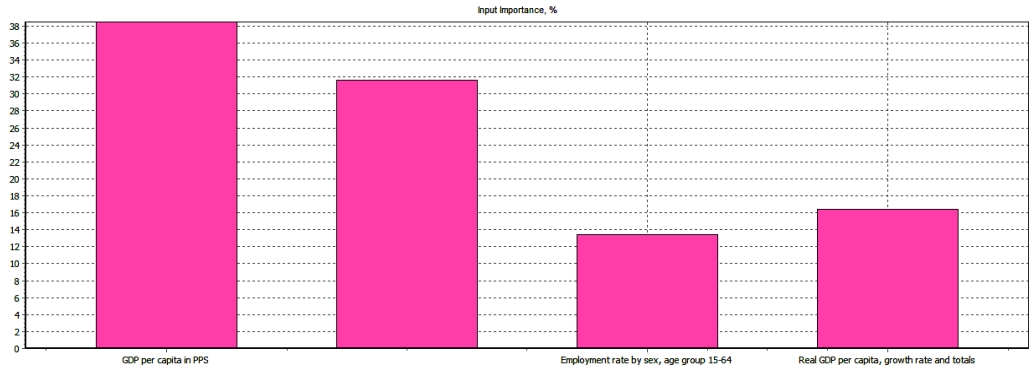
After the training, the ANN evaluated the importance that every input data has over the output data. The result of the evaluation is presented in table 3 and figure 3.

Table 3: Input Data over Output Data Importance

Input column name	Importance, %
GDP	38.48
GOVB	31.66
EMPR	13.41
RGDP	16.43

As can be seen the most influence over the PAS evolution is showed by the GDP, followed by closely by the GOVB. Also, the EMPR and the RGDP have much smaller influence over the evolution of PAS.

Figure 3. Input Data over Output Data Importance



In the next figures, the results of the training are shown:

Figure 4. Evolution of the Network Errors.

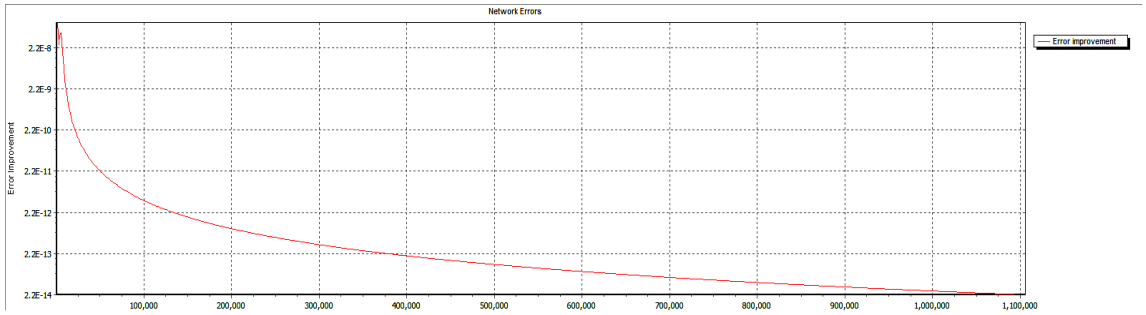


Figure 5. Comparison between Training Set Dataset Errors and Validation Dataset Error

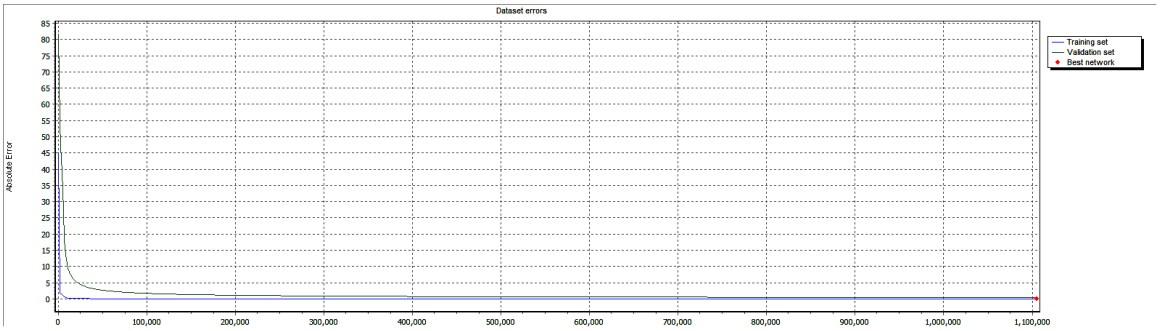


Figure 6. Training Errors Distribution.

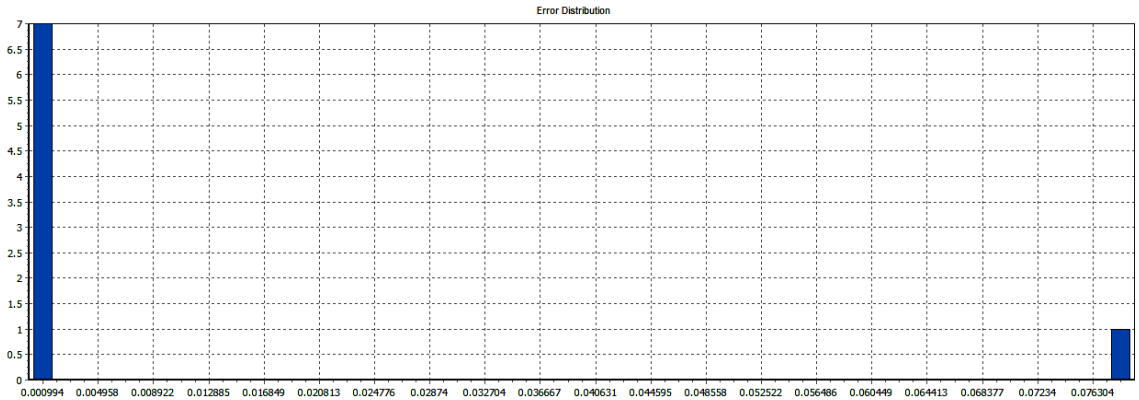
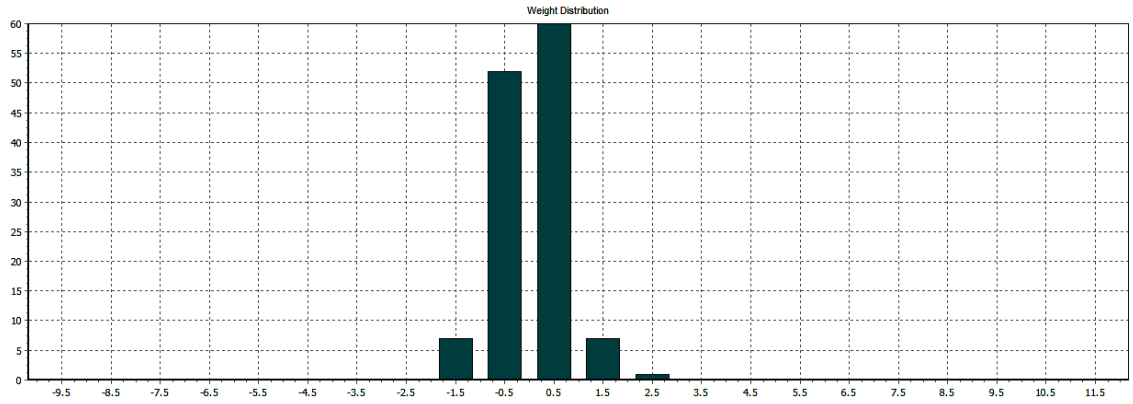


Figure 7. Network's Weight Distribution.



Phase 5 – Validation and Testing.

The last phases of ANN simulation indicate the level of ANN preparedness regarding the expected results. Testing is a process of estimating quality of the trained neural network. During this process, a part of data that was not used during training is presented to the trained network case by case. Then forecasting error is measured in each case and is used as the estimation of network quality.

For the present research, two different sets of data were used for the testing and validation of the ANN training. The first set is the one already used in the training process, as Automat testing (see phase 1). The results of this testing is presented in table 4.

Table 4: Test No. 1. Automat Testing Results – Actual Vs. Output

	Target	Output	AE	ARE
Mean	3921.42	3922.255718	0.920145	0.00023
Std Dev	63.857526	64.983233	2.596286	0.000648
Min	3831	3831.079277	**0.000003**	**0.881**
Max	116	4019.378538	**8.699985**	**0.002172**

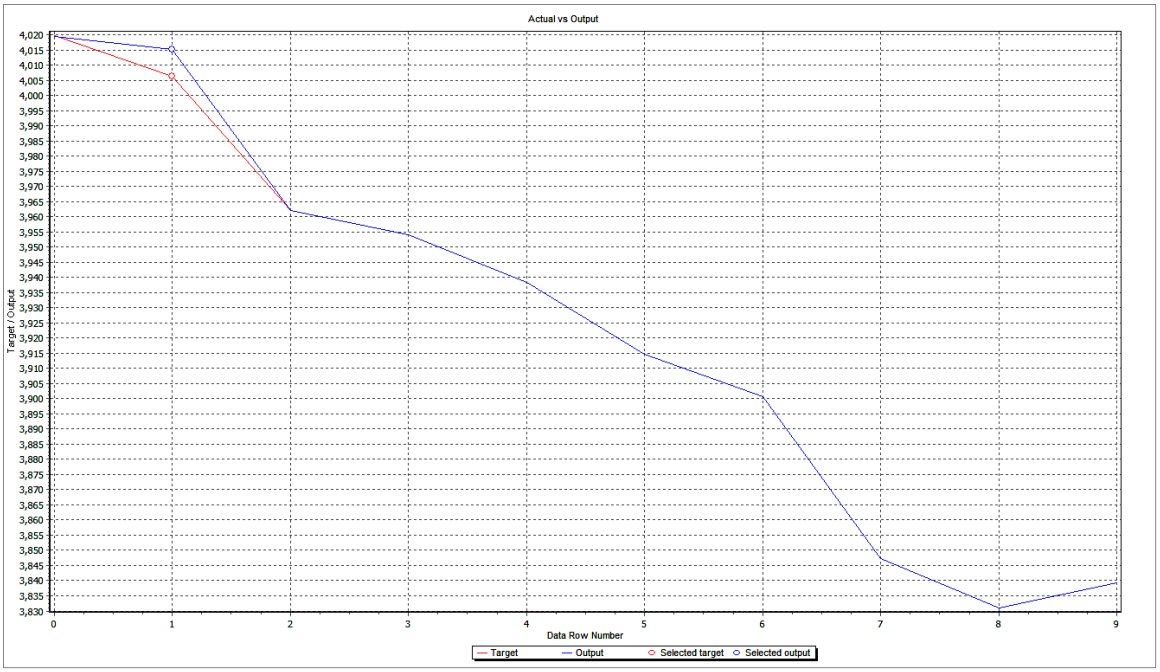
- AE is the absolute error as the difference between the actual value of the target column and the corresponding network output, in absolute values.
- ARE is the absolute relative error as the difference between the actual value of the target column and the corresponding network output, in percentage.

The second set is formed from data that was never fed to the ANN in order to train it. So this set was new to the trained ANN. The results of the comparison between the real data and the simulated data are presented in table 5 and in figure 8.

Table 5: Test no. 2. Testing results - Actual vs. Output

GDB	GOVB	EMPR	RGDP	PAS_real	PAS_simulated	Difference PAS_real vs PAS_simulated [%]
47.00	0.74	59.00	7.50	3903.50	3879.15	0.623799
47.00	0.71	58.60	-6.40	3879.20	4019.27	-3.6108

Figure 8. Actual vs. Output values. Automat testing



From the previous results, the following conclusions are drawn:

- The training phase was successful, the maximum absolute error was 8.699 or the maximum absolute relative error was 0.0022;
- Testing the training with the new data set resulted in a difference between the real data and simulated output, smaller than 3.7%.

4. Acknowledgment and Conclusions

The research established the necessary theoretical characteristics for the pre-processing, training, validating and testing the ANN for the simulation of the numbers of Romanian's Pupils and Students considering the country's indices for economic activity. The ANN training is considered a success and all the initial conditions are respected.

In the beginning of the paper, it is shown how the ANN was trained and tested. The results for the ANN training before the testing were:

- The training error was 8.699 in absolute value and 0.022 in relative value;
- The automat testing deviation of the training was [0.000003; 8.6999] as absolute value and [0.881; 0.002172] as relative value;
- The second testing (with data never feed to the ANN) results produced a relative deviation between [-3.6108; 0.623799];

All the tests confirm good results of the training process as the relative difference between the real ESI and the simulated ESI are smaller than the initial conditions (smaller then 5%).

Also, the ANN offers information about the importance that the input data has over the PAS. The most influential inputs are the GDP (38.48%) and the GOVB (31.66%). This confirms the real situation that the countries with the biggest GDP and GOVB have the most numerous pupils and students.

Overall the ANN training and use is considered a success and the future use of the ANN can be implemented for the simulation of PAS from different Member State countries and even for the entire European-area. Also, the researchers consider that the studies must be extended over a larger period of time and also over a bigger database. This can create a more prepared ANN, and thus a more accurate result.

5. References

- Akira, H. (2003) *Complex-Valued Neural Networks: Theories and Applications*, World Science Publishing Co Pte. Ltd, Singapore, 2003
- Ajith Abraham, Bernard de Baets and all (2006) *Applied Soft Computing Techniques: The Challenge of Complexity*, Springer-Verlag Berlin Heidelberg
- Alyuda Neurointelligence - Http://Www.Alyuda.Com/Neural-Networks-Software.Htm
- Barbeş, L., Neagu, C., Melnic, L., Ilie, C. & Velicu, M. (2009) *The Use of Artificial Neural Network (ANN) for Prediction of Some Airborne Pollutants Concentration in Urban Areas*, Revista De Chimie, Martie, 60, Nr. 3, Bucharest, P. 301-307, ISSN 0034-7752
- De Xiaohui Liu, Paul R. Cohen,Paul Cohen (1997) *Advances in Intelligent Data Analysis, Reasoning about Data*, Second International Symposium, IDA-97, London, UK, August, Proceedings, page 58
- Dumitru, I. N. (2000) *Reţele Neuronale Artificiale. Procesarea Avansată a Datelor (Artificial Neural Network. Advance Data Processing.)*. Printech Pub., Bucharest
- Fahlman, S. E. (1988) *Faster-Learning Variations on Back-Propagation: An Empirical Study*, Proceedings of 1988 Connectionist Models Summer School
- Ilie, C. & Ilie, M. (2010) *Moldova Republic's Gross Domestic Product Prevision Using Artificial Neural Network Techniques*, OVIDIUS University Annals, Economic Sciences Series, Volume X, Issue 1, OVISIUS University Press, P. 667-672, ISSN 1582-9383.
- Ilie, C. & Ilie, M. (2010) *Prevision Research of USA Retail and Food Services Sales Using Artificial Neural Network Techniques*, OVIDIUS University Annals, Economic Sciences Series, Mechanical Engineering Series, Volume 12/No. 1, OVIDIUS University Press, Constanta, ISSN 1224-1776
- WASZCZYSZYN Zenon (2000) *Fundamentals of Artificial Neuronal Networks*, Institute of Computer Methods in Civil Engineering

SELF-HEALING, QUALITY CHARACTERISTIC OF MOBILE APPLICATIONS

IVAN Ion
Professor, Ph.D., Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, Bucharest, Romania, ionivan@ase.ro

BOJA Cătălin
Assistant professor, Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, Bucharest, Romania, catalin.boja@ie.ase.ro

ZAMFIROIU Alin
Ph.D. student, Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, Bucharest, Romania, zamfiroiu@ici.ro

Abstract: *The reliability and security of a software application are two of the most important software quality characteristics because they describe the ability of the software to run without failures and to protect user data. Mobile applications concur with desktop applications in terms of rich interfaces and functionalities and are becoming one the most used type of software applications. Based on the "anytime, anywhere" paradigm, mobile applications must provide special measures to avoid failures and to preserve a high level of reliability and security because mobile operating systems provide limited or none access to administrative or system tools that will allow a user with an IT background to access temporary or persistent data. A high level of software reliability is directly influenced by a low level of failures. We there-fore describe self-healing as a required quality characteristic for mobile applications and we propose a metric for measuring it. This approach is included in the general context of mobile applications quality and the papers describes types of mobile applications, their development cycle and features that are particular to mobile applications quality.*

Key words: mobile, application, self-healing, quality, software metric, regeneration

JEL classification: C2, C88

1. Mobile applications and devices

Mobile applications are software applications developed to run on mobile devices that have resource restrictions, both hardware and software. Based on (Wiki, 2012) and (Schiefer and Decker, 2008) a mobile device is a portable small, hand-held computing device with independent data processing capacity, having a display screen with touch input and/or a miniature keyboard, which can communicate with other IT systems by a wireless connection.

In (Ivan, Zamfiroiu and Milodin, 2011) and (Schiefer and Decker, 2008) mobile devices are divided into distinct categories based on performance, usage, communication capabilities, operating system, input modes, software feature:

- *basic terminals or simple phones* performed basic functions such as send / receive SMS, receive / initiate call; they are not able to use any wireless data communication and multimedia characteristics are limited;
- *feature phones* are mobile low-end mobile devices that have additional functions over and above standard mobile phones like camera, touchscreen and Wi-Fi capabilities; because technology changes rapidly features phones may be considered the low-end smartphones;
- *smartphones* are high-end feature phones with advanced business functions, office applications, high-speed Internet connection, specific multimedia terminals functions, high-resolution touchscreens; they evolved from classic personal digital assistants (PDAs);
- *multimedia terminals* in addition to basic functions present additional functions such a powerful camera, large storage media, audio and video player;
- *fashion terminals* includes performance functions like multimedia terminals and has a special design;
- *mobile Internet devices* represent an intermediary class of mobile devices that are more than a smartphone and less than a tablet PC in terms of features; they are used mostly for Web browsing based on a wireless network connection; modern tablets are becoming the mobile Internet devices of the moment;

- *mobile standard PCs* are mobile computers that have all the features of a personal computer; Tablet-PCs, laptops/notebooks and ultra mobile PCs (UMPC) distinguish from a classic desktop computer by size, display, weight and peripherals.

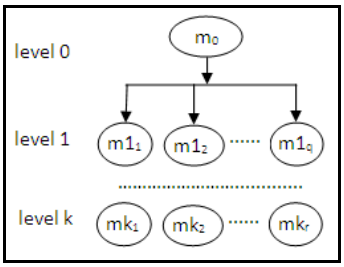
Applications for mobile devices have special features that make them more attractive than the desktop counterparts, but also more sensitive in the development:

- must provide the same user experience across a large diversity of terminals; this requires that the application maintain the same quality level for all types of terminals; in terms of software quality, is important to provide a high level of portability from one terminal to another; applications must adapt and provide the same user experience for any display size;
- users are numerous and heterogeneous in terms of cultural and social background, experience, age; this diversity requires the application to be user-friendly and to be easily accessed by all types of users;
- applications manage personal or other sensitive data that requires high security; mobile devices are more vulnerable to theft or loosing and for that, the information used by mobile applications require a higher degree of security;
- real-time decisions based applications requires accuracy and speed in mobile applications;
- availability on a 24 hours base;
- mobile applications relies on Internet connections for data access, updating sessions and even download; almost all mobile applications are downloaded from Internet based application markets.

These features impose strict requirements on the quality of applications and a high quality of quality is directly related to their degree of fulfillment. In opposition, non-quality leads to catastrophic problems due to failures generated by the mobile applications.

In order to manage this complex construction of quality characteristics, features and restrictions, applications are divided into modules which have a hierarchical structure, as in figure 1.

Figure 1. The hierarchical k-tier structure of the applications



This mode of organization of the applications allows a better management of application code and modules can be developed by different programming teams and later merged to form the entire system.

2. Mobile applications quality

Software quality is an essential concept for software applications because it provides tools to quantify something that for most users is subjective and provides data used to evaluate and compare items in the same category.

Based on the particularities of the mobile environment, a sub set of quality characteristics are considered more important for mobile applications (Ivo, 2005):

- *start time* - a short starting time for the application is very important because most mobile applications are used for activities with short duration and frequency; if the start time is long, the user need to wait too much time for that application; because of the importance of a minimized start time, most mobile applications are not closed but minimized;
- *receptivity* - if the application does not respond in time to restore the user can control, and so there will be two or more tasks to be solved, which will further aggravate the application activity. It is recommended that when an order is received, the user is notified that the order is processed, because they do not give another command;

- *interface* - user interaction with the application should be minimal and the results to be those expected. The application must not contain many modules, modules that will not be used or are used by very few users. An application must be designed for a specific target group and to solve their problems.

In terms of quality characteristics, the set of set of quality factors defined in the ISO 9126-1 (ISO, 2011) quality model can be used to define the mobile application quality:

- *functionality* is a set of functions that satisfy specific application needs in terms of appropriateness, the accuracy, interoperability, security and compliance with standards in the area;
- *reliability* describes the application ability to work correctly in difficult or special conditions. High tolerance to errors and recoverability are two of the requirements that the application in question must have;
- *usability* is defined by the effort required to use the software by users. Operability is high because the time it produces the desired result is very short;
- *efficiency* of a software application is by the correlation optimum between the consumption of resources and complexity or difficulty of the problem to solve. Response time is very important in an application under review;
- *portability* of an application is possibility to use the same application on more devices. If the user change his device with another one who have another operating system is necessary to have his application on this device too.
- *maintainability* is needed effort to make the necessary changes in that software application. The software must be designed as an open product, easily updated and developed by addition, making the product easy to maintain.

Should be considered that the effort to write documented code, orderly and easy to understand and to change is very important to avoid problems that may arise in the future;

- portability is defined as a facility of a software application to be transferred from one medium to another, running on multiple platforms. The application must be easily installed on multiple types of computing machines and can coexist with other programs installed on that machine;
- complexity is a feature on the evaluation of difficulty of developing an program be defined mathematically, linked to the algorithm itself or empirically through volumes of text, different indicators, qualitative.

If classic applications must be equipped with quality features such as reliability, maintainability, portability, as specified, online applications must be friendly, accessible and stable. In contrast to the classic applications, the mobile applications must be robust, correctly, to guide the user, non excess resources consuming, with memory to return back.

3. Self-healing

In many socio-economic fields, mobile software has allowed software developers and users to reach the "anytime, anywhere" paradigm in terms software usability. Because of the extent use of mobile devices and applications, software services are highly available despite the user location. As a result of this special feature, mobile applications are included in socio-economic dependable systems like critical infrastructures, transportation and medical systems. In these systems, high availability, security and reliability are critical characteristics which have to be satisfied all the time. In order to accomplish the requirements regarding availability the applications and the system must be able to self-heal. Based on the concept definition given by (Oreizy et al., 1999), (Keromytis, 2003), (Tichy et al., 2005), (Saha, 2007) and (Salehie and Tahvildari, 2009) the system is able to reconfigure by autonomously redeploying or restarting software components affected by a hardware or software crash failure. Self-Healing is the capability of discovering, diagnosing, and reacting to disruptions (Kephart and Chess, 2003). Accordingly, a self-repair action consists of a set of components scheduled for re-deployment and an appropriate new deployment for each of these components, (Tichy et al., 2005).

Self-healing is only one component from a set of properties like *Self-Configuring*, *Self-Healing*, *Self-Optimizing*, *Self-Protecting*, that defines a self-adaptive software, described in (Salehie and Tahvildari, 2009), which is able to adapt to any situation, to enhance its performance and to protect its data.

For self-healing, the main objective is to maximize the availability, survivability, maintainability, and reliability of the system (Ganek and Corbi, 2003)

Multiple techniques used to implement self-healing capabilities are described by existing research, each with its own advantages and disadvantages:

- the system provides an automatic self-management routine which detects if one of the services crashed at run-time; in this case, it restarts the crashed service taking into account the given deployment restrictions (Tichy et al., 2005);
- the system generates an antibody when is detected an error or an attack, which is distributed to all vulnerable hosts; these hosts can verify if the attack exists and if the antibody can stop it (Brumley, Newsome and Song, 2007);
- information of a site is copied in another place and when the site crashes, all data of this site has a safety copy and must to be able to make an optimal choice and solve the problem by reconfiguration his data (Saha, 2007);
- after detection of an attack or a fault, is invoked a localized recovery mechanism that seeks to recognize and prevent the specific failure in future executions of the application; is verified if the problem has been solved re-running the application against (Keromytis, 2003);
- the application is monitored continuously for failures; after the attack is selected a strategy and the system dynamically modifies the application, using binary injection, so that it is able to detect and recover from the same fault in the future (Sidiroglou, 2009).

In (Dashofy, der Hoek and Taylor, 2002) are presented software architectures for self-healing. For development of self-healing systems two elements are required, an automated decision when to repair the system and the infrastructure of the systems.

The proposed model of self-healing is to develop applications on modules as in (Kephart and Chess, 2003) and when the application is attacked is damaged only few modules which can be very easy restored to initial form. In that way it can be two types of self-healing: horizontal self-healing and vertical self-healing.

Let be an application A, which has the executable source code text stored in a file E and in a file F, which is used. In time due of attacks the F file became F'. If $F' \neq E$ means that there was an attack on the application and the F' file need to be recovered through E file.

If the application is totally in an original state then the application have the self-healing characteristic. Self-healing of an application is the ability to recover from an attack and to restore to the form taken before that attack.

Figure 2. Horizontal self-healing

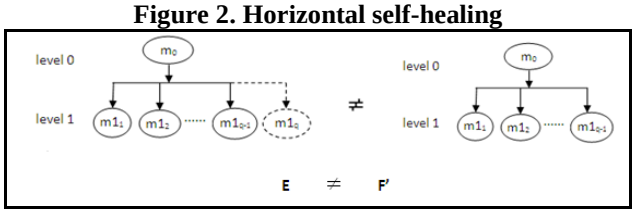


Figure 2 shows the structure of the two applications E and F '. For application F' was attacked and damaged a module on the level 1 of the application. The application F' is different like a structure that the standard application E. Restoring damaged module and bring the application in the form before the attack is considered property of horizontal self-healing.

Figure 3. Vertical self-healing

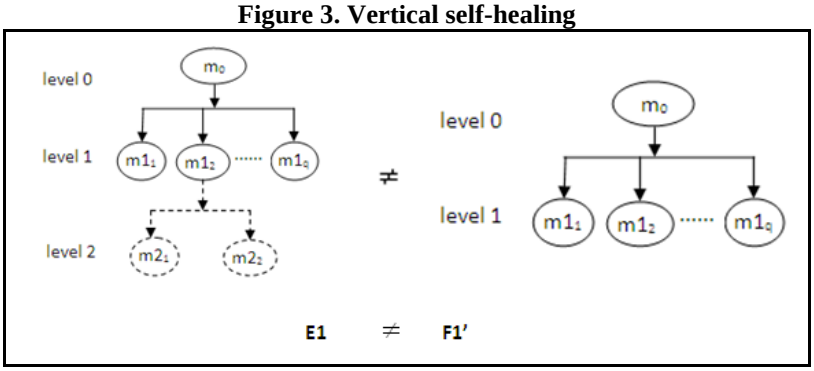


Figure 3 describes the structures of another two applications. For application F1' in an attack were damaged modules from level 2. This application differs from the E1 standard application. Vertical self-healing is the application ability to restore the modules in depth. The self-healing property is specific to open applications that are built as modules interconnected.

4. Software metric for self-healing

Software metric is a mathematical model of indicators developed around an equation, which is designed to measure a characteristic of a software product, taking into account factors influencing the measured characteristic.

For self-healing characteristic is very important period the time that the application F was attacked and became F' and when F' is brought to its original state through the standard application E. In this period there were a number of *ae* accesses the application. Accesses to the fact that the application was attacked are not completed successfully. Thus self-healing indicator suggests I_R is defined by the formula:

$$I_R = 1 - \frac{ae}{at}, \qquad (1)$$

where:
ae – number accesses since the attack until the total regeneration of the application;
at – total number of accesses of the application;
h – number of accesses since the last moment.

$$I_R(ae_0) = 1 - \frac{ae_0}{at} \qquad (2)$$

$$ae_1 = ae_0 + h \qquad (3)$$

$$I_R(ae_1) = 1 - \frac{ae_1}{at} = 1 - \frac{ae_0 + h}{at} = 1 - \frac{ae_0}{at} + \frac{h}{at} = I_R(ae_0) + \frac{h}{at} \qquad (4)$$

Equation (4) highlights the strong relation between variations of I_R values. Thus this indicator is sensitive.

For different values of *ae* the indicator IR have different values, thus this indicator have the compensatory character.

IR have catastrophic character only if the *at* is equal with 0, but this can't be associated with a real scenario.

The indicator value is representative because it shows the speed at which the application is heal and the number of calls the application is minimized destroyed. Formula is applicable to any situations, so the indicator is operational.

5. Conclusions

In our days, mobile applications and devices are complex and for developers is difficult to satisfy the performance requirements required by users and by industry. An efficient approach is to develop applications using hierarchical modules.

The paper presented the self-healing concept as quality characteristic of mobile applications. It has described horizontal self-healing by restoring the modules on the existing level and vertical by restoring the modules on a new level. In order to measure the efficiency of self-healing, a metric is needed to calculate self-healing of applications.

It is recommended that mobile applications to be self-able to update/repair, thus avoiding the inconvenience caused by attacks on them and damage of the source code.

6. Acknowledgement

This work was supported from the European Social Fund through Sectored Operational Programmer Human Resources Development 2007-2013, project number POSDRU/89/1.5/S/59184, "Performance and excellence in postdoctoral research in Romanian economics science domain".

7. References

- Brumley, D., Newsome, J. and Song, D. (2007), Sting: An End-to-End Self-Healing System for Defending against Internet Worms, Sting: An end-to-end self-healing system for defending against internet worms. Book Chapter in *Malware Detection and Defense*, 2007.
- David Garlan and Bradley Schmerl (2002), Model-based adaptation for self-healing systems, In *Proceedings of the first workshop on Self-healing systems* (WOSS '02), David Garlan, Jeff Kramer, and Alexander Wolf (Eds.). ACM, New York, NY, USA, 27-32.
- Eric M. Dashofy, Andre van der Hoek, and Richard N. Taylor (2002), Towards architecture-based self-healing systems. In *Proceedings of the first workshop on Self-healing systems (WOSS '02)*, David Garlan, Jeff Kramer, and Alexander Wolf (Eds.), ACM, New York, NY, USA, 21-26.
- Ganek, A. G. and Corbi, T. A. (2003). The dawning of the autonomic computing era, *IBM Systems Journal, Special Issues on Autonomic Computing*, vol. 42, no. 1, pp. 5-18, 2003.
- Hao Liu K. (2009), A Taxonomy and Business Analysis for Mobile Web Applications, Working Paper CISL# 2009-01, January 2009, Composite Information Systems Laboratory (CISL), Massachusetts Institute of Technology, available online at http://web.mit.edu/smadnick/www/wp/2009-01.pdf
- Hassan Gomaa, Koji Hashimoto, Minseong Kim, Sam Malek, and Daniel A. Menascé (2010), Software adaptation patterns for service-oriented architectures. In *Proceedings of the 2010 ACM Symposium on Applied Computing* (SAC '10). ACM, New York, NY, USA, 462-469.
- Ivan, I.; Milodin, D.; Zamfiroiu, A. (2011) Studierea tipurilor de aplicaţii mobile şi a calităţii acestora. *Revista Română de Informatică şi Automatică*, vol. 21, nr. 2.
- Ivan, I., Zamfiroiu, A. (2011), Quality analysis of mobile application, *Informatică Economică*, vol. 15, nr. 3, 2011, pg 136-152.
- Ivan, I.; Boja, C.; Zamfiroiu, A. (2012) Procese de emulare pentru testarea aplicaţiilor mobile. *Revista Română de Informatică şi Automatică*, pg: 5-16, vol. 22, nr. 1.
- Ivo, S. (2005) *Writing mobile code: Essential Software Engineering for Building Mobile Applications*. Editura Addison-Wesley, Upper Saddle River, 771 p.
- Kephart, J. O. and Chess, D. M. (2003). The vision of autonomic computing, *IEEE Computer*, 36(1), pp. 41-50, 2003.
- Keromytis, A. D. (2003), The Case for Self-Healing Software, Columbia University, 2003.
- Matthias Tichy, Holger Giese, Daniela Schilling, and Wladimir Pauls (2005), Computing optimal self-repair actions: damage minimization versus repair time. *In Proceedings of the 2005 workshop on Architecting dependable systems (WADS '05)*, ACM, New York, NY, USA, 7-6.
- Oreizy, P., Gorlick, M. M., Taylor, R. N., Heimbigner, D., Johnson, G., Medvidovic, N., Quilici, A., Rosenblum, D. S., and Wolf, A. L. (1999), An Architecture-Based Approach to Self-Adaptive Software, *IEEE Intelligent Systems*, 14(3):54–62, May/June 1999.
- Pocatilu, P. (2012) *Programarea Dispozitivelor mobile (in Romanian)*, ASE Publishing House, Bucharest, 600 p.
- Salehie, M. and Tahvildari, L. (2009), Self-adaptive software: Landscape and research challenges. *ACM Trans. Auton. Adapt. Syst.* 4, 2, Article 14 (May 2009), 42 pages.
- Sidiroglou, S., Laadan, O., Perez, C.,Viennot, N., Nieh, J. and Keromytis, A.D. (2009), ASSURE: automatic software self-healing using rescue points. *In Proceedings of the 14th international conference on Architectural support for programming languages and operating systems* (ASPLOS '09). ACM, New York, NY, USA, 37-48.
- Schiefer G.and Decker M. (2008), Taxonomy for Mobile Terminals, A Selective Classification Scheme, *Proceedings of the International Conference on E-Business (ICE-B 08)*, INSTICC Press, July 2008, Porto, Portugal, www.ice-b.org
- StatCounter (2012), Desktop vs. Mobile Internet access between January 2009 – March 2012, available online at http://gs.statcounter.com
- Saha, G. K. (2007), Software - Implemented Self-healing System, CLEI Electronic Journal, volume 10, number 2, paper 5, December 2007.

- ISO/IEC 9126-1 (2001), *ISO/IEC 9126-1 Standard: Software Eng. -Product quality - Part 1: Quality model*, Int. Standard Organization, 2001.
- Wikipedia (2012), Mobile device, available online at http://en.wikipedia.org/wiki/Mobile_device

ECOROLA: A PROPOSAL FOR A VERY HIGH-LEVEL ECONOMICS COMPUTER ROMANIAN LANGUAGE

JOITA Alin-Cristian
Tteaching assistant, Faculty of Computer Science for Business Management, Romanian-American University, Bucharest, Romania, Ph.D. student, The Bucharest Academy of Economic Studies, Romania, alincristianjoita@yahoo.com

BUCUR Cristian
Teaching assistant, Faculty of Economical Sciences, Petroleum-Gas University, Ploiesti, Romani, Ph.D. student ISD ASE Buharest,Romania, cr.bucur@gmail.com

Abstract: *Economic Computer Romanian Language (ECOROLA) is a very high-level domain specific programming language proposal designed for Romanians with an economic background, but with little or no expertise in programming. The main goal is to allow the group of experts from a virtual business incubator to develop applications, with minimum programming skills, for the members enlisted in the incubation program. Proposed model includes both a command line interface and a graphical user interface with a WYSIWYG editor. Results from use of this language and the IDE will materialize in web applications based on HTML, CSS, JAVASCRIPT and PHP.*

Keywords: very high-level programming language, domain specific language, web applications, GUI, graphic-user interface

JEL Classification: C80, L86

1.Introduction

Computers are nowadays present in a myriad of different domains, from managing space programs to managing a personal collection of DVDs. Because of this great variety in the use of computers, since the early 1960, many programming languages emerged, with very different goals. Nonetheless, programming remained inaccessible to many people because of its cryptic nature.

Most computer languages require challenging learning to create structured solutions, have a rigid English-based syntax, harder than natural language and commands that so often seem arbitrary or confusing for beginners or people working outside the field of Computer Science (Kelleher & Pausch, 2005). In addition, the use of a general-purpose language might slow projects down (Bordini, et al., 2006).

After the introduction of computers in all levels of economic processes the researchers looked for a solution to improve the business logic of applications. The possibility of changing application without being a programmer is much efficient. The programmers are required to write applications that can adapt to the business need for a change. One solution to this matter is to have a domain specific language, created for making specific tasks adapted to organization or business domain needs.

In this context, a very high-level programming language is proposed, specific to the field of economics and an Integrated Development Environment (IDE) to enhance productivity by automating tedious coding tasks. Research regarding this type of languages has been conducted since 1975 and created *shell*, *awk* (Bell Laboratories, 1988), *ksh*(Blosky and Korn, 1995), *perl*(Wall and Schwartz, 1990) and many more. A taxonomy is set in (Kelleher & Pausch, 2005) and (Bordini, et al., 2006).

Having less restrictive rules and being easier to understand by an economist, the language can be used to develop tools and modules by nonprogrammers. One of the groups targeted by this model is the team of experts from virtual business incubators, which can enrich its virtual tools, aimed at members enlisted in the incubation program, without even the help of a programmer. That is because after the language script is interpreted, the code is automatically generated in the associated lower-level languages resulting web applications based on HTML, CSS, JavaScript and PHP.

In the same time, according to (Sadat-Mohtasham & Ghorbani, 2008) VHLL are more prone to errors than their lower-level counterparts, unless some influential factors are taken into account, such as: readability, overall simplicity, orthogonality, data types and structures, syntax, writability, reliability and cost.

2.Programming language proposal

Although there are many programming languages today, they are for general utility and not domain-specific. Sometimes using a domain specific language can provide a good expression of specific operations and programming efficiency. At the same time a language oriented to economic domain, with less restrictive rules, and easy to understand for a non-programmer can be very useful to an economist.

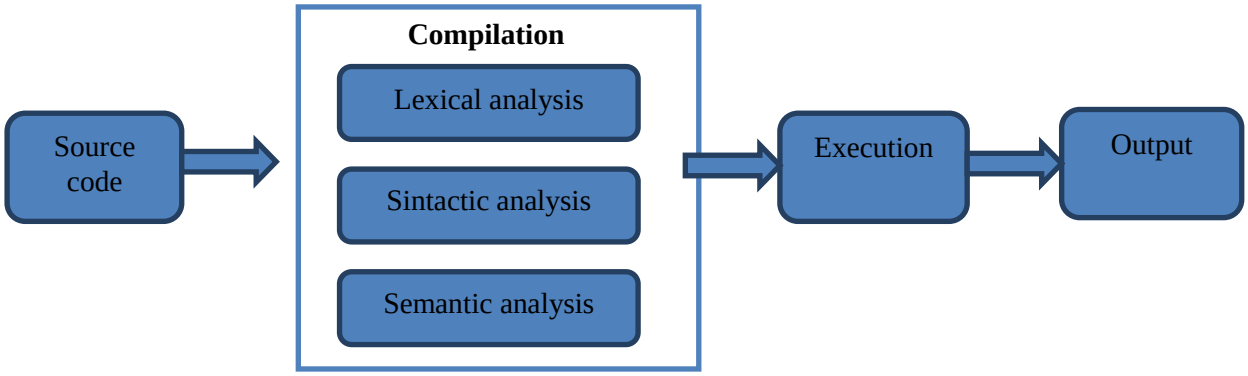
Language sintax

A programming language is a vocabulary and a set of grammatical rules designed to instruct a computer to perform some specific operations (Fields, 2010), (Abderrahim, 2011) and (Plesoianu). Every language has a set of keywords and special syntax for organizing program instructions. Regardless of language used, we must convert the program into machine code, a form of the program that is understandable by machine. This can be done by two ways (Iztok, Iztok, Mernik, & Brest, 2011):

- Compiled languages (eg: C/C++. C#)
- Interpreted languages (Perl, Python, Php)

The proposed new language, for making economic calculations is an interpreted language. It will read text (source code) from the console; it will parse and interpret it (execute instruction) via PHP language.

Figure 1: Main steps in analysis and execution of proposed language

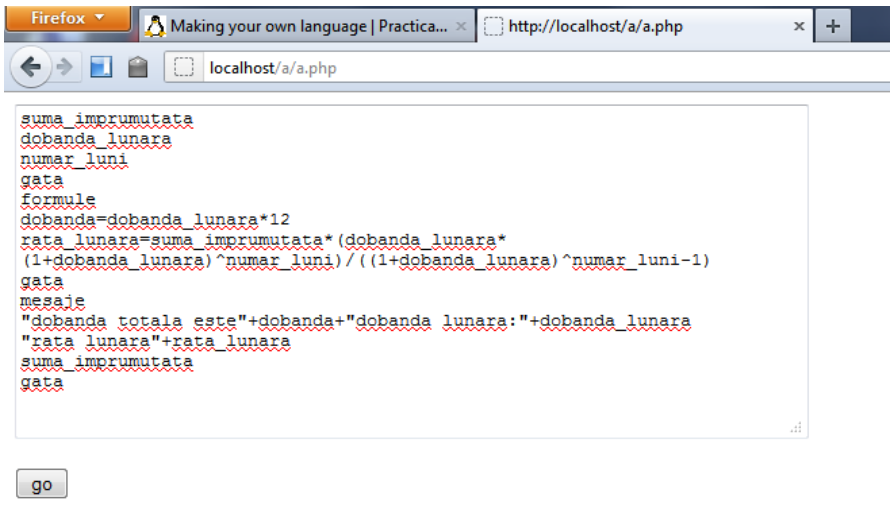


The syntax of proposed language include:

- Keywords: *inputs, formula, posts, ready*
- Operators - known mathematical operators: =,+,-,/,^
- Separators: *(),+,""*

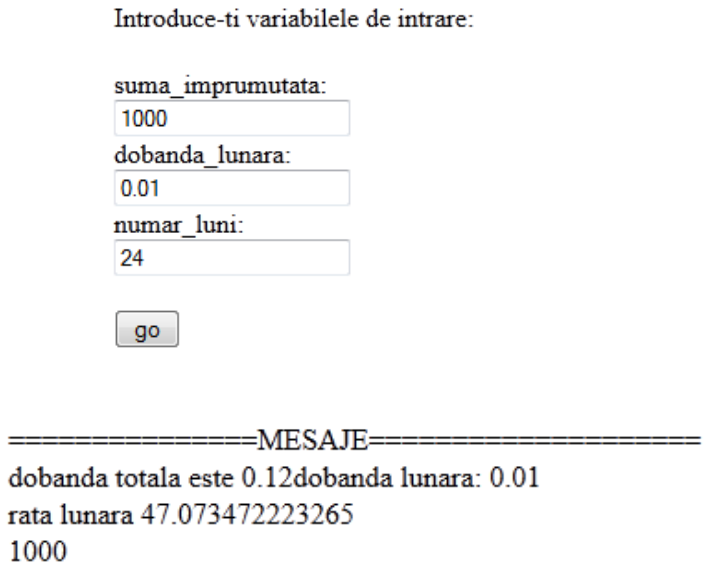
Writing of program by user is done effectively in a browser window, in a text input area <textarea> from an HTML form. The source code will be submitted for analysis and interpretation through execution of form by pressing the GO button.

Figure 2: The console of program



During the program execution, a web form with text boxes it will appear for entering the input variables. After their introduction and submitting of form by Go button operations are executed and the output specified by the programmer will be displayed.

Figure 3: Variables input and output executed by program



Use of language:

Proposed language allows the introduction of instruction in blocks, named using specific keywords, as logical structure of a program. Instructions will be written one on each line, with lower case, without specific sign for termination of the line. We have a block of input "*intrari*", in which the input variables are specified. Variables can be of two types: numeric and string type. The beginning of the block is marked with the keyword "*intrari*" and its completion by keyword "*gata*". In this delimited area will be entered the variable names consisting of alphanumeric characters and character "_", without containing spaces within their names. An example of input block is:

```
intrari
suma_imprumutata
dobanda_lunara
numar_luni
gata
```

The next block is the one for specification of the formulas used in calculation. Block is introduced by the keyword "*formule*" and ends with "*gata*" keyword. Formulas are introduced one on each line using the names of variables, operators of separation () and mathematical operators presented in the described syntax above. As example, having the monthly payment calculation formula, we enter the next formula block:

```
formule
dobanda=dobanda_lunara*12
rata_lunara=suma_imprumutata*(dobanda_lunara*(1+dobanda_lunara)^numar_
luni)/((1+dobanda_lunara)^numar_luni-1)
gata
```

The output block is introduced by the keyword "mesaje"; completion is signaled by the keyword "gata". In this block, the user specifies messages returned by the program and output values. Outgoing messages will be written between quotation marks and for output values will be written the name of the associated variables. On a line of instructions, we may have an outgoing message, a variable or more messages and variable grouped using the separation operator "+".

Language interpretation

The compilation process for a language, according to (Sierra, Fernández-Manjón, & Fernández-Valmayor, 2008), (Limited, 2010) and (Tratt, 2009) can be divided into three parts: analysis, synthesis and output (result). In case of an interpreted language, the synthesis stage, which involves the transformation code in an intermediate language, is missing. The analysis phase involves several phases: lexical analysis, syntactic analysis and semantic analysis.

In the first phase, the source code it is read line by line and it is made the separation of lexical tokens (logical units of language). As the tokens are identified by lexical analyzer they are sent to syntax analyzer (parser) for checking the rules defined by language syntax. After the code has been separated in tokens and parsed, semantic analyzer checks the validity of the script logic.

For parsing code we use three vectors. $ var is an associative array used to store input variables. These variables will be set as keys in array. Formulas are stored in the associative array $ f, where the indexes will be variable names in left part of assignment and right part (actual formula) will be stored as the value of the index. For outgoing messages, will use the vector $out:

```
$var=array();
 $f=array();
$out=array();
```

Passing through each line of source code it is identified the logical block of program it belongs, and appropriate to the block, are extracted the input variables or stated formulas:

```
if ($linie=="intrari"){
      $k++;
      $linie=trim($lines[$k]);
      while($linie!="gata" && $k<$i){
        $var[$linie]=0;
        $k++;
        $linie=trim($lines[$k]);
      }
  }
  if ($linie=="formule"){
      $k++;
      $linie=trim($lines[$k]);
      while($linie!="gata" && $k<$i){
        $formula=explode("=",$linie);
        $f[$formula[0]]=$formula[1];
        $k++;
        $linie=trim($lines[$k]);
      }
  }
```

In messages block instructions are stored in array $out; each line of array $out is an array containing outgoing messages and output variables corresponding to one line of code:

```
if ($linie=="mesaje"){
      $k++;
      $linie=trim($lines[$k]);
      $l=0;
      while($linie!="gata" && $k<$i){
        $iesiri=explode("+",$linie);
        $out[]=$iesiri;
        $l++;
        $k++;
        $linie=trim($lines[$k]);
      }
```

For evaluation of mathematical formulas and expressions we use the class EvalMath (http://www.phpclasses.org/package/2695-PHP-Safely-evaluate-mathematical-expressions.html). It is used a stack type structure in which will be added and removed (pop, push) in a given order parsed operands and operators. Determining the order of operations will be done using the algorithm RPN (Reverse Polish Notation) using a postfix notation in which operators are placed after the two associated operands. During the implementation, each element from the stack is removed, it is checked that it is operand or operator, until they are removed two operands. To them it will be applied the latest removed operand out of the stack and the result is pushed back in stack along with the previous operator. The process is repeated until the stack is empty.

At execution is checked every token of output block, so if it is a message, identified by quotation marks, it is displayed, and if is an output variable, the assessed value is displayed instead of the variable name.

```
foreach ($out as $lout)
  {
   foreach ($lout as $varout)
        if (preg_match('/"([^"]+)"/', $varout, $msg)) {
        echo $msg[1]." ";
        } else {
        $expr="".trim($varout)."";
        $result =  $m->evaluate($expr);
        echo $result;
        }
   echo "<br />";
  }
```

3.IDE proposal

An improvement in modus operandi was included, in the form of an Integrated Development Environment. Such a system may present a syntax editor, graphical tools intended primarily for interface design, support for debugging, compiling and running the program.

This proposal aims to even out the learning curve and decrease design time, creating the interface. For a beginner programmer or a non-programmer is obviously harder to develop an application by writing code than by drag-and-drop operations with existing content in a visual environment (Kats, Kalleberg, & Visser, 2010). In addition, we consider the methodologies for software development uncommon knowledge (Recker, Safrudin, & Rosemann, 2012); to this extent, the user will also need to acquire this information. Therefore, it is necessary to blueprint a wizard module with a graphical user interface of a What-You-See-Is-What-You-Get (WYSIWYG) type. Compared to the command line interface, one can remark certain benefits and disadvantages (Sadat-Mohtasham & Ghorbani, 2008):

+ the learning curve is smaller due to the look and functional similarities with the operating system Window-Icon-Menu-Pointer(WIMP) interface;

+ interface is more intuitive because of the graphical symbols, most operations being self-explanatory. Every control should trigger a tool-tip mechanism to display a short description;

+ complex dependent tasks with multiple steps for application design can be easily grouped in a manner that reduces input errors;

+ ability to embed interactivity (media);

- more computer resources such as memory and processing power must be engaged in order to display a GUI;

- user is limited by the options offered in the interface. Customization of controls is possible only through input fields, so new properties, objects based on inheritance and polymorphism cannot be generated. Nevertheless, this is not an issue, as targeted users are not programming-savvy;

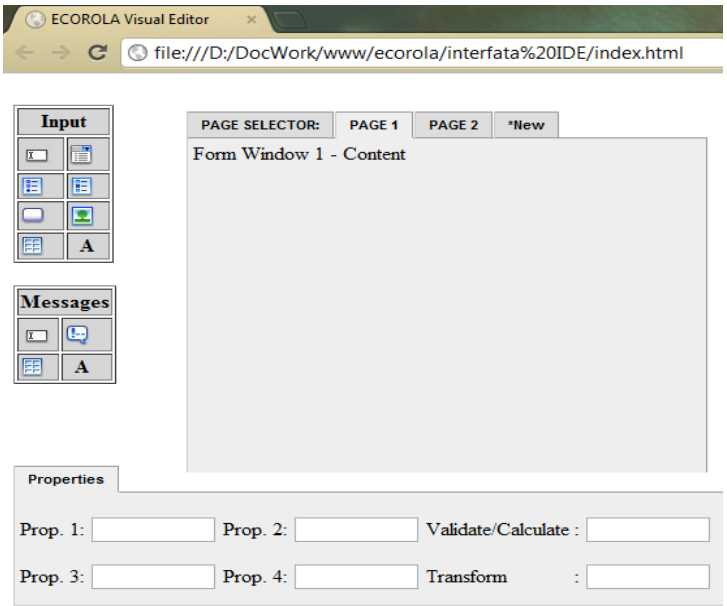
+/- work with mouse and keyboard can be both productive and slow.

As with the command line interface, the GUI will group elements in Inputs, Formulas, Messages and Formatting. Applications are developed in four stages: General Information, Interface design, Testing, Publishing.

General Information requires the user to present the problem and its solution. A title for the application, author information (name, e-mail), short descriptions of the matter and how is solved will be required from the user; date, time and identification data (e.g. IP address) are obtained automatically and stored in hidden variables. In order to publish the application, all fields are mandatory, although, this information will not be displayed in the generated application, unless explicitly requested by the user, in the following stage. After this point, a unique id for the application is generated and data is saved in cookies and on the server. The id may be used later, during development, to track the application. Every section is accompanied by descriptions and help messages.

Interface design stage has a collection of standard elements: Form Window, Page Selector, Toolbox for Inputs, Toolbox for Messages and a Property Bar. Initially, the wizard creates a form page by default. The first step indicated is to insert input fields from Input Toolbox. If more pages are necessary, then additional forms can be created from Page Selector by pressing Add. Every control added to forms can be pressed to reveal and modify its characteristics in Property Bar. Controls are presented in figure 4.

Figure 4: Visual Interface Designer



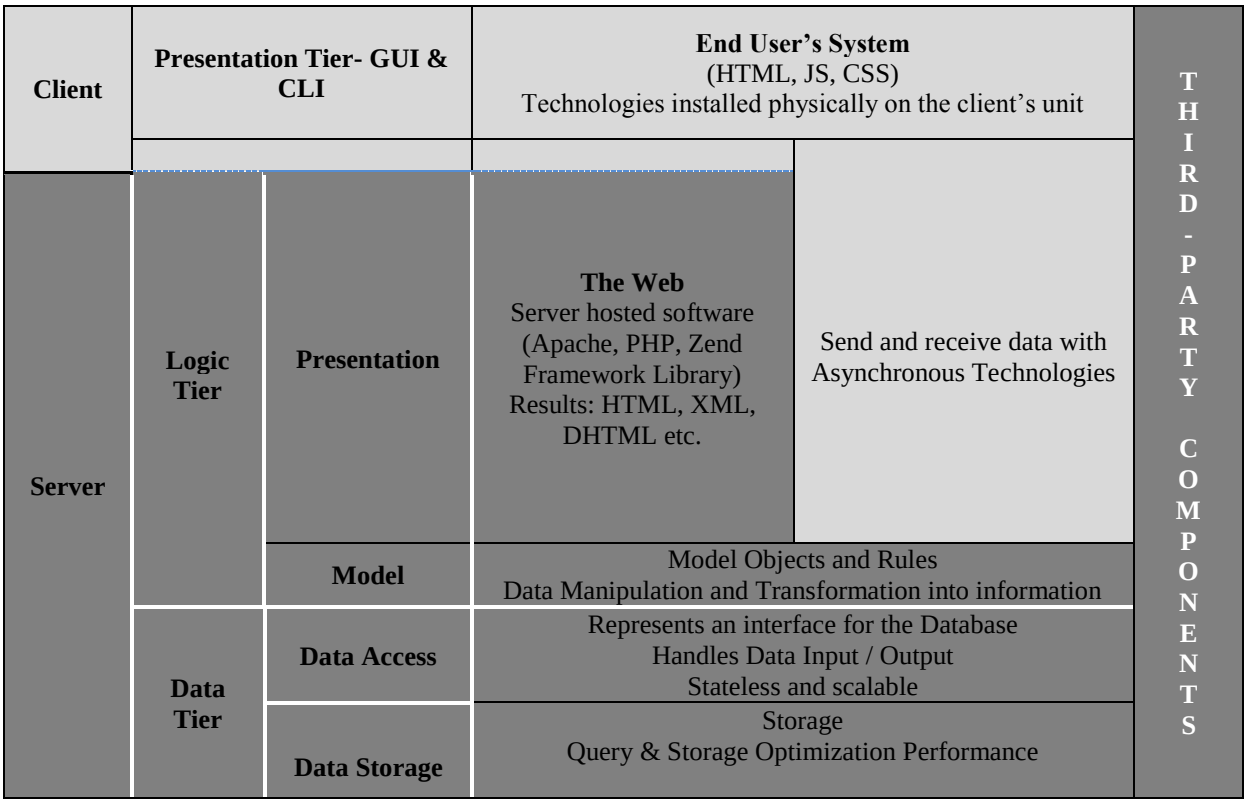
If the input control needs validation or if the inserted data needs processing then validation and transform, fields can be completed with formulas. Formulas may be selected from a list or can be typed. Formula lists are populated from the database.

Before testing phase, all the properties belonging to elements from the interface are saved in the database and translated into ECOROLA. This is necessary as future development might occur in command line interface. In order to test the application, the interpreter can convert the instructions from the proposed language or it can build the interface from the saved objects in the database. The result would be in the form of code, HTML and PHP. If errors occur during building process, description messages are displayed and the wizard reenters Design Stage. If there are no bugs, interpreted code is saved in the database. In addition, the system generates an identification key for the application and all its elements, so editing and deployment can take place.

In publishing stage, the wizard shows the user an HTML code to embed in third-party websites and the application webpage catalog is updated with the new addition. If the author is not a trusted user, the script will not be available to use until approved by appropriate staff members.

Software architecture for the model proposed is client-server three-tier as presented in the **figure 5.** All controls added to the form generate through client-side language JSON elements. These are sent through asynchronous technologies to a server-side script where they are processed and sent forward to be stored in databases.

Figure 5: Client-Server three-tier proposed model



Presented figure includes third party components intended to connect the current system to a virtual business incubator or a similar entity. Architecture for this middleware is not included, more details are available in (A. C. Joita, 2010) and (Căruțașu, 2009).

The proposed model is composed as following:

- Client Tier – consists in the presentation layer that includes the web browser, associated technologies and the transfer platform between client and server. At this segment content loaded in the browser sends requests, receives response packets and using the result generates HTML pages. At the same time, it houses the plugins, so that client-side code can run natively, such as lines of JQuery or JavaScript code.
- Server
    - Presentation Logic Tier – can be considered the Gateway and is implemented as an HTTP server. This segment contains the rules and the methods to data display. When loading a page, the Client layer passes the request to the Presentation Layer on the Server. In this context, header analysis of the request message leads to returning the demanded web page type containing a page template. Objects received from the Logic Tier Model are processed based on formatting rules and used to populate the template with data. API or SOAP requests are issued to 3rd Party services to receive the necessary external objects. References to external libraries being used are added to the template and the final result is sent to the customer. Returned content is text based as HTML, XML or in the form of binary files (images, SWF etc.). At this level, asynchronous technologies such as AJAX are used to avoid loading a new page or to preload information.
    - Model Logic Tier – uses the hardware and software resources of an Application server to implement the logical model. This is the level where code is interpreted and converted to lower or higher level language;

- Data Tier – implementation is supported by a MySQL Database Server.

The multi-tier feature states that the model includes various software packages installed on different servers designed to work together through the use of Web standard technologies (Roşca Ion Gh., 2004).

The prototype, on client-side, implements Presentation tier with jQuery, JavaScript, HTML and ECOROLA. On the Server-side, an Apache server with PHP module assumes the Logic Tier, and a MySQL server represents Data Tier.

4. Conclusion

The process of creating a new programming language involves two stages, the design and implementation. In this paper we presented the stage of design and steps involved in implementation of a domain specific programming language.

ECOROLA is a very high-level programming language designed for economics. Programming environment contains a command line interface and a wizard for a Graphical User Interface. Its goal is to allow economists with little or no programming skills create software that can be embedded in third-party websites through iFrames. Source code is interpreted to lower-level instructions, such as PHP, to generate content in HTML, JavaScript and PHP.

Future research will improve the syntax and the semantics, adding more keywords for describing complex variable types, as arrays or matrices, predefined operations, output formats, and will increase the number of controls and properties available for the visual editor, in order to allow more types of problems to be solved with this language.

5. Aknowledgements

This article is a result of the project POSDRU/88/1.5./S/55287 „Doctoral Programme in Economics at European Knowledge Standards (DOESEC)". This project is co-funded by the European Social Fund through The Sectorial Operational Programme for Human Resources Development 2007-2013, coordinated by The Bucharest Academy of Economic Studies in partnership with West University of Timisoara.

6. References

- C. Joita, G. C. (2010). Technology And Business Incubator Centers - Adding Support To Small And Medium Enterprises In The Information Society. *17th International Economic Conference The Economic World' Destiny: Crisis and Globalization?*, 107-115.
- Abderrahim, A. (2011). *Create your own programming language*. Retrieved April 2012, from The Code Project: http://www.codeproject.com/Articles/50377/Create-Your-Own-Programming-Language
- Bordini, R., Braubach, L., Dastani, M., Seghrouchni, A., Gomez-Sanz, J., Leite, J., et al. (2006). A Survey of Programming Languages and Platforms for Multi-Agent Systems. *Informatica, 30*(1), 33-44.
- Bucur, C., & Tudorica, B. (2011). Solutions for Working with Large Data Volumes in Web Applications. *Conferința Internațională de Informatică Economică - IE 2011 „Education, Research & Business Technologies”.* Bucharest: Printing House ASE.
- Căruțașu, G. C. (2009). Managing Unstructured Data in Implementation of the Integrated Management System (DI - IMS). *16th International Economic Conference*, 10-16.
- Fields, J. (2010). *Business Natural Languages - Domain Specific Languages for empowering subject matter experts*. Retrieved April 2012, from http://bnl.jayfields.com/05_to_web_and_sql.html
- Iztok, F., Iztok, F., Mernik, M., & Brest, J. (2011). Design and implementation of domain-specific language easytime. *Computer Languages, Systems & Structures, 37*(4), 151-167.
- Kats, L. C., Kalleberg, K. T., & Visser, E. (2010). Domain-Specific Languages for Composable Editor Plugins. *Electronic Notes in Theoretical Computer Science, 253*(7), 149-163.
- Kelleher, C., & Pausch, R. (2005). Lowering the barriers to programming: A taxonomy of programming environments and languages for novice programmers. *ACM Computing Surveys, 37*(2), 83-137.

- Limited, F. P. (2010). *Making your own language*. Retrieved April 2012, from TuxRadar: http://www.tuxradar.com/practicalphp/21/5/0
- Patrascu, A., Tnasescu, A., & Bucur, C. (2009). Using Web Technologies in Document Management. *Proceedings of the 13th IBIMA Conference, Knowledge Management and Innovation in Advancing Economies Analyses & Solutions.*
- Plesoianu, F. (n.d.). *Make Your Own Programming Language*. Retrieved April 2012, from http://felix.plesoianu.ro/scratch-lang/index.html
- Recker, J., Safrudin, N., & Rosemann, M. (2012). How novices design business processes. *Information Systems, 37*(6), 557-573.
- ROȘCA I.G., G.-M. B. (2006). *Informatica. Societatea Informațională. E-Serviciile.* București: Editura Economică.
- Roșca Ion Gh., B. C.-S. (2004). *Comerțul electronic - concepte, tehnologii și aplicații.* București: Editura Economică.
- Rotenstreich, S. (1992). Orthogonal languages: A framework for program development. *Computer Languages, 17*(2), 83-100.
- Sadat-Mohtasham, S., & Ghorbani, A. (2008). A language for high-level description of adaptive web systems. *Journal of Systems and Software, 81*(7), 1196-1217.
- Sierra, J.-L., Fernández-Manjón, B., & Fernández-Valmayor, A. (2008). A language-driven approach for the design of interactive applications. *Interacting with Computers, 20*(1), 112-127.
- Tratt, L. (2009). Chapter 5 Dynamically Typed Languages. *Advances in Computers, Elsevier, 77*, 149-184.

MODERN COMMUNICATION MANAGEMENT USING IT&C STRATEGY

MAREŞ Valerica
Assistant Professor, PhD, The Bucharest University of Economic Studies, Faculty of Accounting and Management Information Systems, Romania, maresvalerica@yahoo.com

MAREŞ Marius Daniel
Professor, PhD, SPIRU HARET University, Faculty of Management, Finance and Accounting, Bucharest, Romania, maresdaniel@yahoo.com

ŢARCĂ Naiana
Associate Professor, PhD, University of Oradea, Faculty of Economic Studies, Romania, ntarca@yahoo.com

Abstract: *Businesses are based on communication, which allows a "whole" to function. The "whole" can be an individual, a team, a community, a nation, a system, the global population. Those that master communication, to perfection, can become "winners" – in the sense that they can look confidently towards the future and can be those who manifest themselves which make them continuously succeed. Communication can be a problem of trust and acceptance or rejection of other people's ideas and feelings and it requires abilities, and developing abilities requires practice. Supervising the general tendencies inside and outside an organization is achieved through communication in order to be sure that the strategy, the vision, the systems are aligned with the realities and in order to anticipate and prepare the desired changes.*

Key words: communication, abilities, supervising tendencies, strategy, excellence in business

JEL classification: M15, M21

1. Introduction

The basic elements of the communication process are the *transmitter* or the source (the person that initiates the communication and transmits the message) and the *receiver* or the destination (the person that receives the message). Between the two points of the communication is the *message* (the information) which is codified in a signal. Between the transmitter and the receiver takes place an exchange of information, or message communications. Messages can get transmitted using the *verbal* language (using words) or *nonverbal* language (using the body, the space, the time and the things) and the *paraverbal* language, which is a vocal form of the nonverbal language (for example the tonality or the inflection of the voice, the rhythm of speech, the way of stressing words, breaks between words, verbal tics, etc.)

A communication (information and message) has to be transmitted in such a way that the receiver can understand it, receive it, record it and accept it. The answer of the receiver of the transmitted message proves that the message was well understood. His reaction is called feedback and closes the communication circle, because the receiver, at his turn, codifies information (the answer to the message) and communicates it to the transmitter. Thus we can consider that in the communication process the roles are continuously changed: the receiver becomes transmitter and vice versa.

In order to be efficient you have to obey the three basic rules of feedback:

- ✓ **Immediate feedback** means that as soon as you understood or you think you understood the message, but after you paraphrased and asked the clarification questions;
- ✓ **Honest feedback** means a discontent or a powerful frustration;
- ✓ **Supportive feedback** means being gentle or plainly civilized, saying what you have to say without hurting or destructing. Moreover, in a moment of tension, you can transmit in this way the care and the consideration for the other person. Supportive feedback shows that two communicators consider the problem equally important.

There are numerous differences between "saying" and "communicating" or between "hearing" and "listening". "Saying" is a one way process, and "communication" supposes a two-way communication process.

The channel is the physical means of transmitting the message, also called the hypothetical "way" or the "path" followed by the message which in the modern era takes various forms such as:

- Technological channels: telephones, tape recorders, computers, videos, pagers, radios, iPod's, CD-players etc.
- Written channels: mails, reports, files, memos, forms, books, magazines, newspapers;
- Face to face channels: conversations, interviews, meetings, presentations, trainings, and lectures.

The technical or physical means that transforms the message in signal is called *media*, and the communication media can be oral or written, depending on the communication means that we use; talk-listen-observe, thus we communicate in the oral media, or write-read, so we use the written communication. *The technical support* of the communication channels comprises the technical means that can support the communication process, such as: telephone, computer, fax, telex, email, portable, tablet, internet and audio-video means.

In the communication process can interfere perturbing elements, which take the generic name of "noise" and are responsible for distorting the message (noises, some smell, an inappropriate physical state, etc.). Feedback measures the success of the communication process.

2. Interpersonal communication and the IT&C means

Interpersonal communication is the most important and the most used communication form. People can't avoid this type of communication because their social existence depends on the ability with which they can initiate discussions with others. The family life, the relations with friends, the professional activity, all depend on this quality.

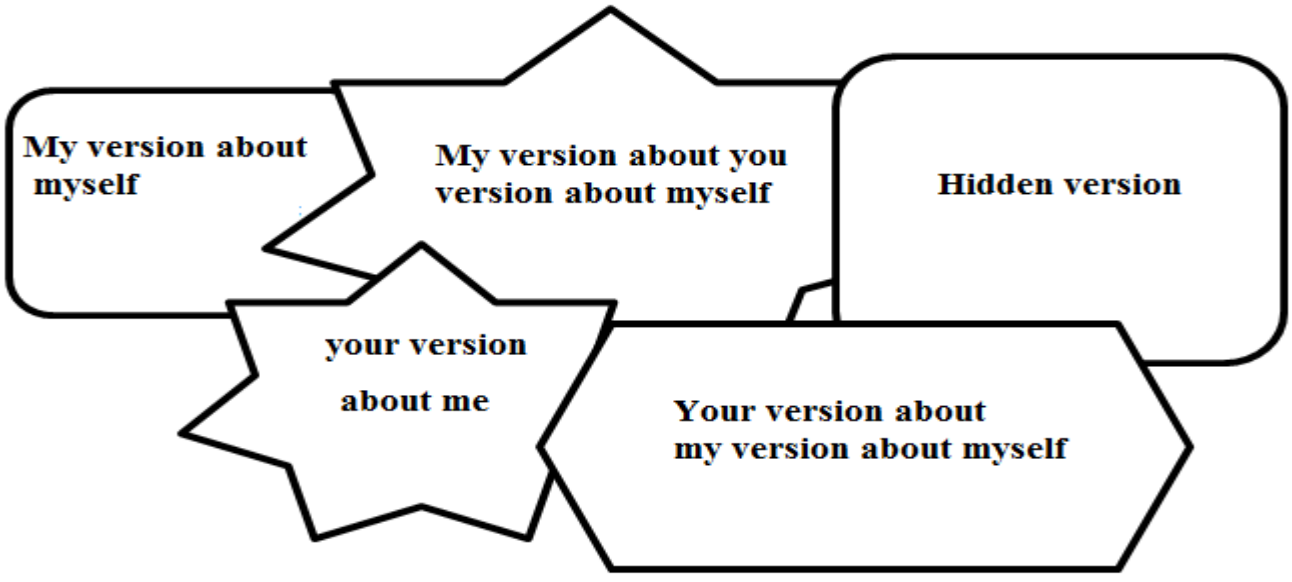
Interpersonal communication refers to face-to-face communication. This type of communication is important in order to understand yourself and to build your relationships with others. In order to achieve this one has to develop his capacity of auto-analysis, self-knowledge, self-exposure as well as the barriers and the perturbing factors that make the communication process more difficult.

Freud considers that knowing yourself, being completely honest with yourself, is a significant human effort because this self-honesty demands searching, discovering and accepting information about self and a self-improvement desire.

Psychoanalysts consider that the first law of life is not biologic preservation but preserving the image about yourself, which is a basic human need. The image, the opinion about yourself, which comes from organizing our experiences, perceptions, values and objectives depends on our "reality" and determines our behaviour. When we communicate, we talk to the image we have about our interlocutor, and he will communicate with us according to his version about us and about himself.

Any person have at least five images over himself, each of them permanently influencing itself and changing as a consequence of the communication process.

Figure 1: Images over our self



The harmony between these versions make communication simple, direct and no distorted. This harmony can be achieved by "testing" the image about yourself (self-knowledge) or by opening to others,

self-exposure, "reflecting ourselves in others". We are social beings, and our personality results from asociation not from isolation.

During our life we have to face urgencies and importance's when it comes to taking decisions but one of these factors tends to dominate. The option for urgency or importance has a profound effect on the results that we obtain. The dependence on urgency is an auto-destructive behaviour which satisfies temporarily a hole created in unfulfilled necessities. When the urgency becomes the dominating factor in our life, the sense of importance is lost. More and more people are becoming to discover and solve cronical problems that hinder them from dealing with the things that are really important for them. As we go from urgency to importance, we often rise the question "What are actually the priorities and how can be put them on the forefront of our life?" According to Stephen Covey there are three fundamental ideas that lead to the answer to this question:

a) Fulfilling the human necessities and capacities is covered in the sintagm "to live, to love, to learn, to leave something behind" – which represents our physical, social, mental and spiritual needs.

b) The truth of the principle of the "*authentically north*", which offers context and significance to the place in which we are or where we want to go, the way in which we get there – based on the eternal law of cause and effect which functions in the world of human efficiency and interdependence. During our lives and in society we pick, unrelently, what we sowed. Searching for "shortcuts and miracle solutions" is based on the illusion of rapid solutions, and this affects not only the conscience of our fundamental needs, but also the way in which we try to satisfy them.

c) The potentiality of the four human qualities:

- self-knowledge – the power to detach from ourselves and to examine the way in which we think, the motivations, the past, the actions, the habits and the tendencies
- consciousness – our internal system of ethical and moral guidance, which allows us to realize when we act or to contemplate ourselves while we act, to connect to the wisdom of the old ages and of the soul, to become aware of the feeling of our unique features and mission.
- independent willpower – our capacity to act with consciousness, morality and visionaries, to act based on principles, not on emotions and circumstances, to be the product of our choices, to be responsible (that is to choose independently of our temporary dispositions and tendencies)
- creative imagination – the power to forecast a future situation, to conceive a plan and to solve the problems synergically. It allows us to make a personal engagement, to establish a purpose, to plan a meeting, to look at ourselves as we fun fill the commitment even in the most difficult conditions.

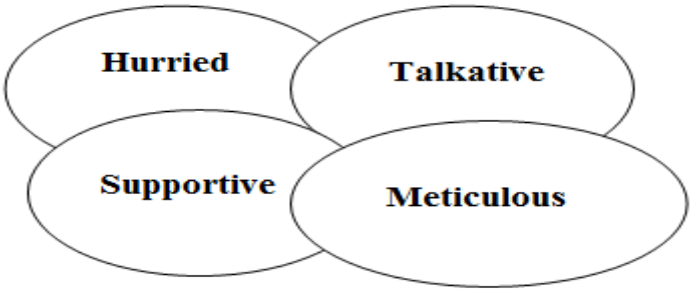
These form the basis of our decisions, of human freedom which manifest itself through the power to make choices, to react, to change. There is some space between stimulus and reactions. Perfecting the four qualities underlines above supposes a process of exerting with a continuous interest, our modern IT&C means such as: forums, chats, emails, sms's, socializing pages and sites (Facebook, LinkedIn etc.).

3. Efficient communication using IT&C means

Efficient communication requires abilities, and their development requires practice. While wishing to express the content as well as the intent, we learn the language of logic and of emotion – this last one being by far the most powerful and with the biggest motivating effect. We listen first with the eyes and the heart, then with the ears. We try to understand the intention of the communication without prejudice. We invest time and patience, trying first to understand and expressing our feelings honestly, we get to present our point of view by starting to demonstrate a clear understanding of the other points of view. Initiatives are concentrated on creating a spirit of cooperation and by improving the communication process, especially when it comes to preparing people for the ability to listen and explaining clearly their point of view.

In order to set a diagnosis, in any situation, we can start by understanding first the requirements, the pursuits, the forces and tensions and the situation of the partners in order to offer solutions meant to satisfy the requirements and the shortcomings, to creating bridges between the necessities of the people involved and the services – the offered products and then collecting facts in order to understand the situation. Identifying the type of client (see the next figure) supposes different approaches for achieving an efficient communication.

Figure 2 Type of clients



*The **hurried*** is the type that thinks as follows: "I want to solve it now, it doesn't matter how" and has the following features: he's noisy and is perceived as belligerent in difficult situations and makes pressures for obtaining rapidly results. The hurried loses quickly focus and interest listening to vast technical explanations. The hurried persons are competitive, provocative beings, they're always ready for action and are interested in the final result, thus he detests to listen to details and technical explanations and to justify why a service doesn't function. In order to communicate with such persons we have to be direct and to the point, not loose ourselves in details and justifications and to keep a powerful, clear, energetic and confident stance, maintaining a quick rhythm of speech and to give him the impression that we are rapidly approaching (and we really have to lead him) a concrete result, giving the impression that we make things move and we act decisively in his favour. While communicating with such a person we have to use words like: *quickly, immediately, now, today, result.*

The **talkative** is social and has a tendency towards interaction with others, having a need to interact with people and to achieve recognition, being an optimistically, inspired, convincing and confident person, who talks with enthusiasm, smiles most of the times and expects from his communication partner a similar behaviour. Apart from having their requests solved, the talkative clients expect that those they're "serving" should be friendly and open. The talkative people tend to laugh and to speak continuously and you will see them talking all sorts of things not because they are absent minded but because they love conversation more than the subject. They are those who run into people on the street, being attentive to certain people and things. In order to communicate with this type of client you have to be smiling and energetic, enthusiast, friendly words and tone of the voice can conquer en enchant him. When we talk we have to be as creative as possible, describing the situations in lively, colourful images and the rhythm in which we talk should be quick and lively. In a communication with a talkative we'll have to use often words like: *attractive, I feel that, imagine that, extraordinary*.

The **supportive** is interested in the way things are going. It is a reserved person, who doesn't like change; he is patient, relaxed, logical, systematic and flexible. When he sees an inconvenience because the offered product doesn't work, he will first make sure that he doesn't bother you with his complaint because he has a strong need for stability.

We recognize easily the supportive, because they are loyal to the company in which they work, they often wear clothes or accessories with the mark of the company where they work, they accept verified solutions, avoiding risks and unusual situations and are the type of client that once conquered, if the service supplier doesn't make big mistakes, can keep for a long time. For a communication with this type of client the voice should be warm, calm and confident, the voice should be quiet and the rhythm of the voice should be relaxed and thoughtful and the undertaken actions should be rhythmic and relaxed. In a communication with a supportive we will meet more often words like: *promise, security, trust, security*.

The **meticulous** is the type of client interested in the quality of services. He should be entrusted that in any moment everything works perfect and things are done "a la carte" following the rules and procedures precisely. The meticulous is a perfectionist and wants things to be done precisely, in order and with accuracy. He needs facts, graphics and data in order to declare himself content and follows precisely the rules and procedures in every situation. The meticulous client keeps the same appearance in different situations, he doesn't want to show what hides under this and doesn't agree easygoing and free discussions when he calls, he will ask for elaborate solutions, and he adores to "breaking the thread in four". Communicating with a meticulous client is a true art. They are very suspicious regarding what we want to tell them. Thus for an efficient communication we have to permanently keep a controlled tone of voice, with few modulations, we have to be direct and precise in what we say. The rhythm has to be slow

and thoughtful and the words, the tone of voice and the rhythm has to be adapted accordingly. In a communication with a meticulous we will meet words as: *demonstrated, informed, analyze, think some more*.

Report is the name given to the process in which two people have a good relationship. It means building trust, harmony and cooperation in the relationship with the other. The key to developing a rapport is acceptance.

4. Elements that target excellence in communication management

When we communicate, we connect through consciousness (listening and answering, acting in harmony with our interior voice) to wizdom and become a person with character and consciousness. The quality of this connection can increase through several means:

- Reading and meditating over wisdom;
- detaching ourselves and learning from our own experience;
- noticing carefully the experience of those surrounding us;
- finding time to calm down and listen to our interior voice;
- answering to our internal voice;
- developing independent willpower, taking and following commitments.

Excellence is the attribute of those educated and brave and the elements of excellence in managerial communication are:

- *Personnel*. The personnel are the supreme value because its members are those that program and achieve everything.
- *Self*. Change and improvement has to start with ourselves, with things that we can control directly
- *Style*. The participative style generates innovation, initiative and devotion as well as a lot of unpredictable behaviour
- *Skills*. Skill that comprise delegation, communication, negotiation and self-management are fundamental for achieving high performances and can be learned and developed by a continuous learning and training.
- *Shared vision and principles*. An agreement over fulfilling tasks, from which everybody wins, in which the parts that have the same vision based on the same principles gives freedom to the parts to do what they wanted to achieve.
- *Structure and systems*. We have interdependency relationships with many people, and this interaction requires a structure and some systems in an equilibrium state – at high levels of productivity.
- *Information system* about the risk factors, in order to take wise decisions.
- *Efficient system for financial and psychical compensation*, which rewards the synergic cooperation and creates a team spirit.
- *Programs for an efficient preparation and development of personality,* in which the preparation is, controlled less by the system and more by the "pupil", which means he is free to progress in his own rhythm and to choose the methods for attaining the goal agreed by both parties. Recruiting and selecting people in the interest of all parts, according to the work experience of the individual and with the way of attaining success which the activity demands. The job description should contain as many from the people's interests and skills, and should state the new growth possibilities, leaving a certain degree of autonomy for choosing the methods for obtaining the desired results.
- *Communication.* Face to face meetings between the manager and the employee for defining the win-in agreement over task fulfillment, together with the meetings with the personnel, with agendas oriented towards action, a system for keeping a track on the suggestions coming from the employees and a way to reward the ideas that lead to savings, policies and procedures of open doors, annual interviews for quick promotions, studies based on anonymous opinions or ad-hoc brainstorming meetings, are the key to an efficient communication in an organization – especially if they are organized based on a shared goal and vision
- *The strategy* that reflects the declared goal, the available resources and the state of the market, frequently renewed to reflect the changing of the external conditions.

Supervising the general tendencies inside and outside an organization, in order to make sure that the strategy, vision and systems are aligned with the exterior realities and to anticipate and prepare the desired changes.

5. Excellence in business communication

An efficient communication requires expressing the content and the intention, speaking in the language of logic and of emotion, the latest having also the biggest motivating effect. Communication is first of all a problem of trust and of acceptance and rejection of the ideas and feelings of the others. If we succeed to leave apart the adversity spirit, the social norms, the exaggerated attention for building our own image, we save a lot of time and energy.

At the root of many communication problems are the problems of perception and credibility. Each sees the world through the prism of his own reference systems – influenced by feelings, convictions and behavior. Most credibility problems can be solved if at least one of the involved parts realizes that the issue is a perception problem.

Communication is a process that requires the skill of using empathy, in which reality is tested, in which individuals try to clarify the expectations in a relationship and separate the people from problems, concentrating their attention over the common interests, enumerating options from which all those involved can gain something, insisting over the use of objective criteria.

Some people use frequently the tactics of empathic listening, as:

- imitating the content (the least efficient);
- reformulating the content;
- sentimental reflection;
- reformulating the content and sentimental reflection.

These tactics are efficient only if they derive from an honest desire to understand. Learning the empathic way of communication transcedes the techniques and requires a lot of time, time that is later saved. There are several levels of answers, such as:

- answers in an "autobiographical" manner;
- evaluation (we are for or against);
- examination (we ask questions from a subjective perspective);
- we give advice starting from our own experience;
- we impersonate (we try to represent people, to explain their motivation, their behaviour, based on our own motivations and behaviours).

During the communication process one can distinguish **useful attitudes** such as:

- good will and lack of distrust in the honesty or the mental equilibrium;
- the desire to solve the perceptive differences;
- Lack of resistance to the opinions of others and the preparation for change.

Moreover choosing useful behaviours such as

- listen in order to understand;
- talk so that you can be understood;
- starting the dialogue from a common reference point or from a point over which you agree and approach gradually the disagreement subjects is a solid base for idea exchange and creating an easy communication.

In the daily life as well as in the virtual communication medium the usual listening levels are:

- ignoring the peer by refuse or Cancel;
- pretending we listen while we perform several tasks simultaneously;
- selective listening on certain themes;
- careful listening (active, reflexive);
- empathic listening represents a way of listening with the intention to understand (first listen, in order to truly understand) what is being said which means an as deep as possible understanding, deeper at intellectual and affective level.

This supposes more than just recording, reflecting or understanding the words that are being said (experts say that 7% of communication is done through *words*, 38% through *voice tonality* and 55% through *body language*). In order to be a good speaker it is first of all required to have a good capacity of expressing ideas. It supposes a simple expression, so that the interlocutor can understand accordingly the transmitted message and a correct pronunciation of words can be easily recognized and the message is transmitted in a positive form. The *voice tonality* is important on the phone where gestures can't be seen

and that is why it is more difficult to communicate over the phone, where we can't make use of gestures and face expressions, but only of voice. The voice stamp expresses feelings and emotions through:

- ✓ **Inflections** – that underline words and syllables for intensifying the message;
- ✓ **Pitch** – how high or low frequency the sound is;
- ✓ **Rhythm** – how many words you speak in a unit of time;
- ✓ **Volume** – how loud or soft the voice is heard.

- *Empathic listening* supposes perceiving feelings, significations, listening to the compartmental language – using both cerebral hemispheres, receiving messages from the intimate of the interlocutor, understanding the premises of the partners. Empathic listening supposes as well taking risks, you become vulnerable and in order to exert an influence you have to let yourself be influenced.

Essentially, after we understand our communication partner we try to be understood. In order to make an efficient presentation we communicate empathically with his intellect, knowing that he knows that we understood his intentions, objectives, worries. We set the accent on the human element, in the same degree as we stress the financial or the technical element – by widely opening the door to creative solutions, discovering new alternatives, developing the interpersonal relationship by accumulating trust.

6. Conclusions

Verbal or nonverbal communications are very important from a social perspective. Usually the verbal statements are influenced by a series of factors such as: the fear of not offending or upsetting, the desire to close a deal, social pressure which sometimes makes us say that we agree with someone even if it is not true etc. The nonverbal communication completes, strengthens, emphasizes the sense of the verbal messages and in certain situations, it is even more credible than the verbal communication. Although in some situations the nonverbal communication is much more ambiguous than the verbal communication, it is more credible when its signals contradict the signals of the verbal communication.

1) Some people are mainly oriented to the exterior world and they are called extroverts, and some are mainly oriented towards their interior world and they are called introverts. Extroverts are open persons, social, communicative, optimistic, peaceful, benevolent, they agree or they argue with those around them, but they keep relationships with them. Introverts are closed temperaments, hard to penetrate, shy, little communicative, with a tendency towards being dreamy and to have difficulty adapting.
2) A good communicator has to master the verbal communication as well as the non-verbal communication especially when the communication is done through IT&C means which can memorize the images and the voice.

7. References

- Alain, B. (2000) *Cum dialogăm şi cum convingem*, Polirom
- Cave, c. Can computer help zou think? www.members.optusnet.com.au/charles57/creative/software/essay.htm Software for Creativity & Idea Generation – an overview of whatţs available.
- Samuel, P. (2007) *Improving teamwork with Kirton Adaption-Innovation Inventory*, Breakthrough Management Group
- Ghiga, G. (1999) Elemente fatice ale comunicării în română vorbită, Bucharest: Alcris Publishing House.
- Spelber, D. and Wilson, D. (1986) *Relevance: Communication and Cognition*, Oxford
- Stancu Ş. (2001) *Relaţii publice şi comunicare*, Teora
- Stephen R. C. (2000) *Etica liderului eficient sau Conducerea bazată pe principii*, Editura ALFA
- Ventola, E. (1987) *The structure of social Interaction*, London.

THE MARK OF UNEXPECTED FINANCIAL CRISIS LEFT UPON CLOUD COMPUTING

MARIAN Matei
Associate professor, Ph.D. student, Faculty of Economic and Business Administration/Department Business Information Systems, "Al. I. Cuza" University, Iaşi, Romania, mari_martinik@yahoo.com

Abstract: *Governments from all around the world spend considerable amounts on IT, with the aim to improve efficiency, transparency, and to deliver more interactive services to citizens and businesses. Cloud computing is a fundamentally new approach to IT, which may lead to significant cost savings. At a time of financial crisis and of major fiscal constraints, cloud computing may be key to contain expenditures. Cloud computing offers the possibility of cutting costs, while still increasing computing capacity and introducing new products and services with a rapid time-to-market cycle.*

Key words: cloud computing, financial crisis, virtualization, scalability

JEL classification: G01, L86, O33

1. Introduction

"The rise of the cloud is more than just another platform shift that gets geeks excited. It will undoubtedly transform the IT industry, but it will also profoundly change the way people work and companies operate." (The Economist, 2008)

Grappling with the financial crisis, developing and transition countries governments are being forced to make painful choices about budget cuts all while citizens and businesses are demanding more and better services, transparency, and accountability. But today's cost reductions cannot be a one-off tactics – they must be driven by transformational improvements that enable sustainable cost efficiencies for the long-term. A crisis of this scale requires a brand new approach to managing IT investments. This time of financial and economic crisis offers a great opportunity to pause and rethink traditional approaches and start making a transition to the next generation models of IT applications and infrastructure delivery, to get more results for less money. It may potentially offer an opportunity for developing countries to leapfrog to the cutting edge of IT solutions at a fraction of the cost of what it took developed countries to get there. More and more government organizations are starting to explore the Cloud Computing model as one of such transformational approaches. Cloud computing approaches can increase efficiencies and yield dramatic savings in IT costs which can be repurposed to client-facing programs and services.

2. Cloud Computing and The Financial Crisis

Cloud computing may likely be the next foregone conclusion, driven primarily by two key forces: (1) a flexible pay-as-you-need operational cost model and (2) the growth of software-as-a-service (SaaS) solutions and application offerings. If needed improvements in security and performance monitoring come as promised, it will sway CIOs to let go of their data centers and shift to the cloud paradigm. (Hugos, 2011, p. ii)

The "cloud" in the term *cloud computing* is a metaphor for computing resources (hardware and software) that companies and users access without the need to know exactly where that hardware and software is physically located. (Hugos, 2011, p. 43)

Cloud computing is a new outsourcing model for deployment of ICT solutions. It entails use over the Internet of computing hardware and software infrastructure and applications that are remotely hosted and managed by a service provider. Under the "software-as-a-service" (SaaS) model, client organizations pay a fee for use of the ICT solution and need not invest in the infrastructure nor have the skills to manage it. The main notion is to use only what you need and pay for only what you use. Because the cloud infrastructure is regularly upgraded, users always have access to the latest hardware and software without having to upgrade or patch legacy systems.

The World Bank has lent more than $7.73 billion for ICT solutions in more than 100 countries. Despite successes, however, ICT projects have run into difficulties as the resulting technology solutions are either not successfully installed, are underutilized, are not sustainable or include high opportunity

costs due to delayed implementation. The total cost of ownership of these ICT solutions has been a serious issue for many developing countries. They require major investments in hardware infrastructure as well as sophisticated skills to manage this infrastructure. Licensing costs for the requisite software platforms are very high both initially and throughout the lifecycle of the systems. Cloud computing may offer a radical solution to some of these challenges. (FCCC, 2009)

The cloud model offers a much cheaper way for businesses to acquire and use IT -- in an economic downturn, the appeal of that cost advantage will be greatly magnified. This advantage is especially important for small and medium businesses, a sector that will be a key target in any plan for recovery.

The current economic crisis in the United States may have a silver lining for IT companies that invest in cloud computing, as it will contribute to significant growth in this sector over the next years, according to research firm IDC. Based on a survey of IT executives, CIOs and other business leaders, IDC said that it expects spending on IT cloud services to reach US $42 billion in 2012, a growth of threefold that will in part be bolstered by the current economic crisis that began in the U.S. and spread all around the world. (Montalbano, 2008)

A somewhat neglected aspect of the current financial crisis is the huge spike in trading volumes in recent days. In some cases, they have raised to more than double the average of the months (and years) that preceded the crisis.

Consider a recent Tabb Group report estimating that a mere 1 millisecond delay in response to a trading request can cost a brokerage firm up to $4 million. Or a Google study that an additional 500 millisecond delay in returning a "search results" page decreases user traffic by 20 percent. (Perry, 2008)

Systems, of course, don't care if stocks are going up or down; they just need to handle the transactions. With the increasing possibility of weakening global economies -- and shrinking IT budgets to follow -- companies can no longer engage in the expensive practice of over-provisioning. On the other hand, with increased competition and the increased mission-criticality of IT systems to businesses and consumers, failure is not an option. But failure is not relegated to complete down-time. A system's inability to scale might just result in a slow-down, but this, too, can be extremely damaging to the business.

Scaling operations up and down smoothly as conditions change and the ability to pivot quickly and address new threats or opportunities are what make companies successful in today's economy. Business models with high fixed costs are much riskier than they used to be. What happens if actual demand for a company's products is less than predictions? Can the company still be profitable and cover operating costs if only 60 percent of its capacity is utilized? What if only 40 percent of its capacity is activated? Companies with high fixed cost investments in unused production capacity are risking their profits and their very existence. (Hugos, 2011, p. 23)

Although recent surveys don't show a significant drop in the growth of IT budgets for 2009-2010, many of these surveys were conducted before the recent financial fallout, and its full impact might not be reflected in them. In any case, it is clear that one of the top priorities of CIOs is to cut costs. This is especially true in industries such as financial services, automotive and airlines that were hard-hit by the current crisis and high oil prices.

Cloud computing offers the possibility of cutting costs, while still increasing computing capacity and introducing new products and services with a rapid time-to-market cycle. Companies developing virtualization technologies already have started this trend of server consolidation. Nonetheless, virtualization has only scratched the surface of opportunity. According to Gartner in October 2009, 18 percent of workloads were on virtual servers. That said, Gartner expects that number to jump to almost 50 percent by 2012, impacted substantially by small and medium-sized businesses that are now taking advantage of lower price points. (Messmer, 2009)

Cloud computing is the next step in this evolution. Almost every industry can benefit from such low-cost and elastic compute capacity. With more rigorous analysis and automation, cloud computing can enable whole industries to move from faith-based to fact-based decision-making. A pay-as-you-go model for resource use reshapes the fundamental cost structure of building and operating applications. The initial barrier to starting a project is drastically reduced; and until there is dramatic uptake in the use of an application that has been developed, the costs for running it remain low.

The good news is that this isn't the only cost advantage. By harnessing the cloud, you can also take advantage of cloud providers' economic leverage because of the volume at which they can purchase hardware, power, and bandwidth resources. Cloud computing has the ability to change the economics of

IT. It changes the ratio between capital expenses (CAPEX) and operational expenses (OPEX) because of its pay-only-for-what-you-use principles.

CAPEX are expenses that create future benefits. When a business spends money to buy fixed assets, it incurs a capital expenditure. Accounting adds a capital expenditure to an asset account (capitalizing it). Typically, CAPEX requires a large up-front investment to be amortized (written down) as its value decreases over time. The initial CAPEX purchase can require a large cash outlay - one reason startups are avoiding these purchases altogether. OPEX is an ongoing cost for running a product, business, or system. It's a day-to-day expense akin to costs, such as sales and marketing, and as such, it can much more easily be increased or decreased as business needs dictate. This isn't possible with CAPEX because you've already purchased the equipment. (Rosenberg, 2011, p. 51)

Cloud computing arises from the combination of technologies that have been developing over the last several decades. And the ongoing rapid evolution of cloud technology is driven by the pressing needs of organizations to cope with change in their markets and change in their financial situations. In a time where information and communication technology is now mission critical to every facet of business operations and where "safe bets" are hard to find, it is safer to explore new markets and new ventures on a pay-as-you-go basis instead of investing a large sum of money up front and hoping the investment pays off. Cloud computing makes this possible. It can be quickly rolled out; it can be quickly scaled up to handle increased volumes if business takes off; and it can be just as quickly discontinued or scaled back to cut costs if business does not take off. This variable cost operating model allows companies to replace capital expenses with operating expenses, and that is critical to any organization operating in high-change, unpredictable environments. Cloud computing enables companies to best align operating expenses with revenue and protect their cash flow and operating profits. In addition to its financial impact, cloud computing also affects how companies structure their organizations, how they manage and coordinate their daily operations, and how they engage and motivate their people and their business partners. (Hugos, 2011, p. xiii)

In frequent talks and writings about cloud computing, a recurring question is coming up front: "How do we deploy our existing and new applications in a cloud environment?" Are there any unique technologies that need to be applied to take advantage of the on-demand power of cloud computing? The answer is "it depends".

Mostly stateless applications, such as simple Web applications, can be quite easily ported to cloud environment. After all, these are simply standard servers that are accessed via the Internet. But more sophisticated applications, such as the data-intensive VAR analysis ones used on Wall Street, as well as various transactional Web applications, require some adjustments. Traditional applications servers, such as JEE app servers from IBM, Oracle and JBoss, won't cut in a cloud environment. They tend to have central bottlenecks and are rigidly "wired" to the physical servers. Such applications require elastic application servers that can go, grow and shrink on demand, handle fault-tolerance transparently and be easily moved around without any downtime. Fortunately, there is a new generation of application servers that fits exactly those requirements.

There is another aspect of on-demand computing that is the key to its success: both the hardware and the software need to be provided in a "pay-by-use" based-model. Typically, this would be some sort of CPU hourly pricing for computation power, data size pricing for storage and memory and so on.

According to analysts, average datacenter utilization rates are at the 15-20 percent range. That means that on average, 80-85 percent of the organization's IT resources -- servers, storage, networks and software -- are idle. This creates a huge financial drain on the company. Secondly, predicting demand is an extremely tricky business. The busiest minute last year is not necessarily a solid indicator of the volume during the busiest minute this year. (Perry, 2008)

"I'll never buy another server again" said the Director of IT for a medium-sized Software-as-a-Service (SaaS) company, after recently completing the deployment of a new corporate website for his organization. This website (a PHP based application with a MySQL backend) showcased the corporate brand and the primary online lead-generation capability for the company's business. Before the overhaul, it was run from a redundant pair of web servers hosted by one of the leading managed-hosting service providers at a total cost of roughly $2,200/month. The company replaced the infrastructure for the original website with a cloud implementation consisting of a pair of virtual server instances running for roughly $250/month—almost a 90 percent savings! Its quality of service (QoS) team monitored the performance and availability of the website before and after the change and saw no measureable difference in the service quality delivered to end users. Buoyed by the success with this initial project,

this organization is looking at all future initiatives for the possibility of deployment within the cloud, including a software-build system and offsite backup. (Rosenberg, 2011, p. 6)

According to the article entitled "The Long Nimbus" published by the Economist magazine about the impact of cloud computing on company organization structures, "Businesses are becoming more like the technology itself: more adaptable, more interwoven and more specialized. These developments may not be new, but cloud computing will speed them up." (Siegele, 2008)

3. Conclusions

The attractiveness of this arrangement among adopters is why International Data Corporation (IDC) predicts that IT spending on cloud IT services will be $44.2 billion in 2013 - a significant rise from an already substantial $17.3 billion in 2009. (Gens, 2009)

Internet-based technology is driving economic change at a level not seen since the spread of industrial technology in the late nineteenth and early twentieth centuries. The spread of cloud computing is an excellent example of the phenomenon known as "creative destruction", which was popularized by the economist Joseph Schumpeter. Schumpeter pointed out that in capitalist economies, there are waves of change where the introduction of a new technology or new process for doing things upsets and replaces the previously dominant technology along with the people and companies who used that technology. Cloud computing is having this effect on vendors who sell traditional versions of computing technology and on the people who make their living operating traditional computing technology.

Companies that have large investments in traditional in-house computing technology will not abandon those investments immediately, nor should they. The transition of companies to cloud-based technology will be quicker for some and slower for others depending on their individual circumstances. But the change will happen. History shows over and over again that resistance to the spread of new technologies is almost always futile, and often fatal. People and companies that resist are finally forced out of business and replaced by others that do adopt new technology. Clearly the best strategy for people and companies is to actively explore the opportunities for cloud computing and begin appropriate projects to gain experience in its use and to understand its strengths and weaknesses. (Schumpeter, 1942).

4. References

- FCCC (2009) A Government Transformation Initiative Workshop, Financial Crisis and Cloud Computing: Delivering More for Less Demystifying Cloud Computing as Enabler of Government Transformation.
- Gens, F. (2009) IDC Exchange, Defining ''Cloud Services''–an IDC update. Available at http://blogs.idc.com/ie/?p=422.
- Hugos, M. (2011) Business in the Cloud - WHAT EVERY BUSINESS NEEDS TO KNOW ABOUT CLOUD COMPUTING, John Wiley & Sons, pp. ii, xiii, 23, 43.
- Messmer, E. (2009) Gartner: Server virtualization now at 18% of server workload. Available at http://www.networkworld.com/news/2009/102009-gartner-server-virtualization.html.
- Montalbano, E. (2008) IDC: Economic Crisis to Bolster Cloud Computing. Available at http://www.cio.com/article/print/455568.
- Perry (2008) Cloud Computing and the Financial Crisis. Available at http://www.hpcinthecloud.com/hpccloud/2008-09-29/cloud_computing_and_the_financial_crisis.html.
- Rosenberg, J., Mateos, A. (2011) The Cloud at Your Service - The when, how, and why of enterprise cloud computing, Manning, pp. 6, 51.
- Schumpeter, A., J. (1942) Capitalism, Socialism and Democracy, New York: Harper, 1975, pp. 82-85.
- Siegele, L. (2008) Economist, The Long Nimbus. Available at http://www.economist.com/node/12411864?subjectid=348909&storyid=E1TNQTTRGQ.
- The Economist (2008) Let it Rise. Available at http://www.economist.com/node/12411882.

MOBILE ANALYTICS - MODERN DECISION SUPPORT SYSTEMS

MUNTEANU Adrian
Ph. D. Candidate, Academy of Economic Studies / Department of Economic Informatics, Bucharest, Romania, adrianm21@yahoo.com

PUICAN Florina
Ph. D. Candidate, Academy of Economic Studies / Department of Economic Informatics, Bucharest, Romania, puicanflorina@yahoo.com

Abstract: *Nowadays, company executives require relevant and accurate business information in real-time to take strategic decisions. It is also essential to have freedom to access this information anywhere and anytime. There is a real need to extend this functionality beyond the office and on the fingertips of the decision makers. Mobile tools for Business Intelligence aim to provide these features in a flexible and cost-effective manner. This paper describes a model of architecture of mobile tools for accessing decision support systems to overcome the limitations of existing mobile business intelligence tools and highlights the benefits of using mobile business applications.*

Key words: mobile intellgience, business intelligence, decision support system

JEL classification: D81, M15, M21

1. Introduction

In the modern corporate era, information has become the most important resource for a business enterprise and therefore corporations spend a substantial amount of capital to maintain enterprise systems that record, process and present data from all aspects of the organization. The objective behind the utilization of enterprise systems is to monitor, control and manipulate business processes based on accurate hard facts. The mobile Business Intelligence tools have been envisioned to facilitate the process of decision making through the knowledge presented by such systems. The aim of the system is to provide an open source solution which incorporates flexibility, device independence and cross platform integration together as a single solution to utilize business intelligence capabilities. The need for these features combined into a mobile tool arose from the gap present between the source of data and the decision maker Airinei (2010). Such tools make this process easy, economically feasible and customizable. This paper presents a brief review of the Business Intelligence technology used in the market. It relates this technology to the value of a tool such as the mobile ones to ease of the access to the information in real time.

Having as main purpose to support and improve business decisions and co-operate competitiveness, IT systems are used to collect and process business data. These business intelligence systems (BI systems) strive to combine formerly spread and fragmented data from different parts of a company. Through analysis and transformation, data is turned into information, a basis for business strategic decisions. The increased availability and system performance of mobile systems allow flexible on site data collection and processing, thus extending business intelligence to mobile business intelligence. At the same time different wireless communication properties and heterogeneous system environments must be considered. Furthermore, the use of BI tools on mobile devices (e.g., mobile phones, notebooks, smart phones, tablets), requires to securely adapt to different physical and networking environments due to a changing work context.

2. Business Intelligence architectural components

The early days of business intelligence processing (any variety except data mining) had a strong, two-tier, first-generation client/server flavor. (Some business intelligence environments that were hosted on a mainframe and did querying and reporting were built with a centralized architecture). Conceptually, early business intelligence architectures made sense, considering the state of the art for distributed computing technology (what really worked, rather than today's Internet, share-everything-on-a-Web-page generation). Many of these early environments had a number of deficiencies, however, because tools

worked only on a client desktop, such as Microsoft Windows, and therefore didn't allow for easy deployment of solutions across a broad range of users. Additionally, long-running reports and complex queries often bottlenecked regular work processes because they gobbled up your personal computer's memory or disk space.

Most, if not all, tools were designed and built as fast clients — meaning most of their functionality was stored in and processed on the PC. In addition to the bottleneck problem, all users' PCs had to be updated because software changes and upgrades were often complex and problematic, especially in large user bases.

The "Business Intelligence" term is being used in the modern business world to fulfil the emerging need of analyzing data from large and variable data sources and this need brought an evolution and a new line of competition between vendors of different enterprise systems. The dramatic trendy evolution in IT field from the last decade and the increased demand of data driven knowledge, large companies have enabled themselves to store an enormous amount of data. However, the request to derive information quickly and to facilitate the decision making process of business organizations, a new set of tools was required, which leaded to Business Intelligence. BI means the idea of collecting, storing, analyzing and providing access to data and it consists of a set of applications and technologies that support decision making, querying and reporting, online analytical processing, statistical analysis, forecasting and data mining.

This requirement for Business Intelligence applications and technologies has led to the development of BI software. Most major software vendors have launched their own Business Intelligence modules that can be and are packaged with their software products. Nowadays organizations using enterprise systems are also using the services of available BI software on the market. The functionality of this software is independent of the rest of the modules, but also provides means to seamlessly integrate it with enterprise systems. Business Intelligence systems encompass a database system(s) and various information systems to provide a unified view of data and have to prove able to analyze data and provide information to the end user to make quick and critical decisions Czernicki (2010). In the context of this research existing BI tools and modules have been surveyed from two perspectives: usage and industry. The usage describes the general use of these systems to determine how the system integrates itself with the enterprise system. Later has been determined the existing expectations of the market from BI modules based on the upraising business needs and more complex IT environments.

In order to identify the BI components, we first need to determine what kind of problems are the BI tools designed to solve:

- Structuring data for analysis;
- Analyzing data;
- Creating reports;
- Viewing and sharing reports.

Based on these generic problems identified above, the BI tools should comprise of following important components:

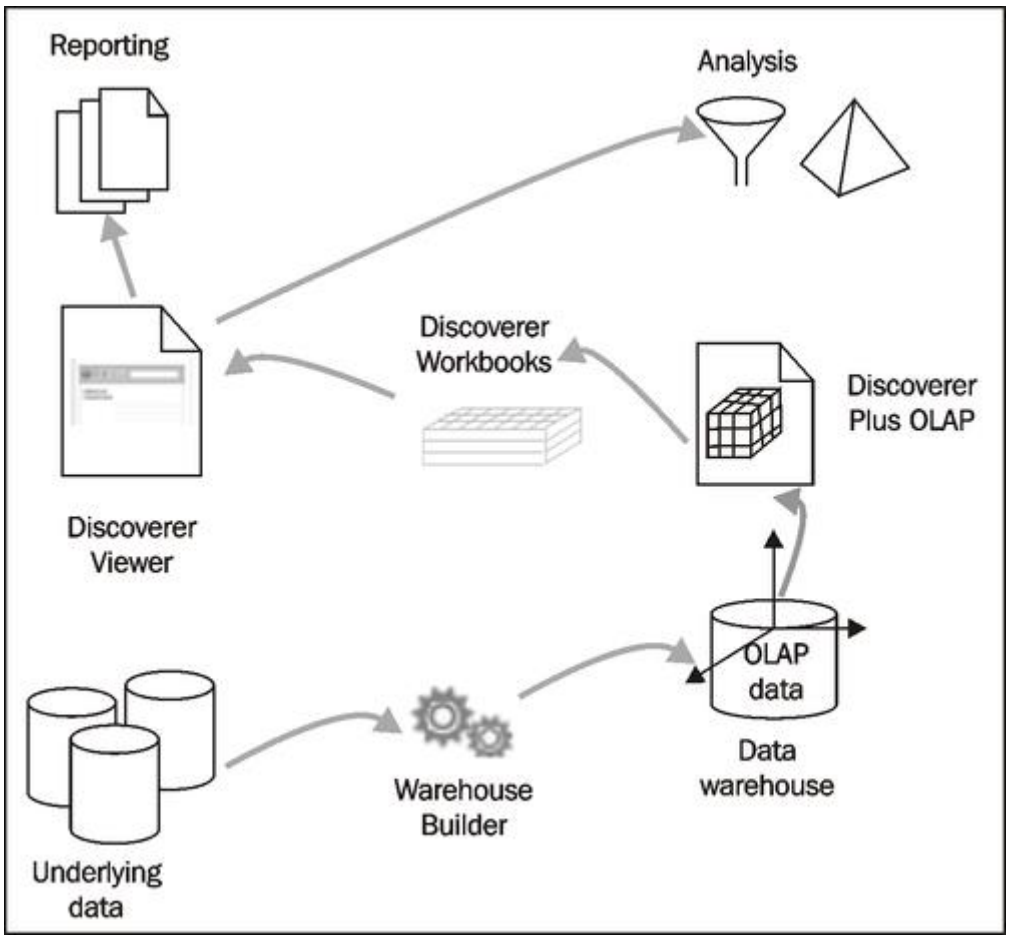
- a component for analyzing relational data and creating reports;
- an OLAP tool for analyzing multidimensional data and creating reports;
- a viewer for reporting and publishing the results;
- portlet creator for publishing existing reports to company Portal;

In addition, there are some useful tools in the usual Business Intelligence tools packages like:

- administrative tools for structuring data for analysis;
- a tool for designing and deploying data warehouses, data marts, and Business Intelligence applications;
- sometimes there is the need for analyzing relational data and creating reports;
- a reports developer for building and publishing reports;
- add-In packages and APIs for sharing data to transform OLAP data in more spreaded formats like MS Excel.

As described above, some components are required to prepare data for analysis, while others, are designed for analysis, reporting, and sharing. Utilizing this concept of complementary components is precisely the method that should be employed when composing a Business Intelligence system by grouping complementary Business Intelligence components together to implement a workflow that suits business requirements best Bâră (2008). Diagrammatically, this might look like in Figure 1:

Figure 1: Example of information workflow between various BI tools.



Source: Laursen (2010).

Concluding, we may say that the beginning of a new era of Business Intelligence architecture has arrived, regardless of whether some tool of choice is a basic querying and reporting product, a business analysis/OLAP product, a dashboard or scorecard system, or a data mining capability. Although BI product architecture varies between products, when one evaluates products that might provide business intelligence functionality for business data warehouse should keep an eye on some major trends:

- **Server-based functionality**: Rather than have most or all of the data manipulation performed on users' desktops, server-based software (known as a report server) handles most of these tasks after receiving a request from a user's desktop tool. After the task is completed, the result is made available to the user, either directly (a report is passed back to the client, for example) or by posting the result on the company intranet.
- **Web-enabled functionality**: Almost every leading tool manufacturer has delivered Web-enabled functionality in its products. Although product capabilities vary, most products post widely used reports on a company intranet, rather than send e-mail copies to everyone on a distribution list.
- **Agent technology**: In a growing trend, intelligent agents are used as part of a business intelligence environment. An intelligent agent might detect a major change in a key indicator, for example, or detect the presence of new data and then alert the user that he or she should check out the new information.
- **Real-time intelligence**: Accessing real-time, or almost real-time, information for business intelligence (rather than having to wait for traditional batch processes) is becoming more commonplace. In these situations, an application must be capable of "pushing" information, as opposed to the traditional method of "pulling" the data through a report or query. Like with traditional data-extraction services, business intelligence tools must detect when new data is

pushed into its environment and, if necessary, update measures and indicators that are already on a user's screen. (In most of today's business intelligence tools, on-screen results are "frozen" until the user requests new data by issuing a new query or otherwise explicitly changing what appears on the screen).

- **Support for mobile users**: Many users who are relatively mobile (users who spend most of their time out of the office and use laptops or mobile devices, such as a Blackberry, to access office-based computing resources) have to perform business intelligence functions when they're out of the office. In one model, mobile users can dial in or otherwise connect to a report server or an OLAP server, receive a download of the most recent data, and then (after detaching and working elsewhere) work with and manipulate that data in a standalone, disconnected manner. In another model, mobile users can leverage Wi-Fi network connectivity or data networks, such as the Blackberry network, to run business intelligence reports and analytics that they have on the company intranet on their mobile device.

3. Mobile Intelligence

Demand for mobile intelligence is exploding, due largely to the popularity of the Apple iPad and other tablet devices. The popularity of mobile BI is also seen by many analysts as a factor that will drive greater adoption of BI.

Mobile BI solutions provide comprehensive support for analytics on smart phones and tablet devices so that today's mobile workforce can access the information they need, wherever they are, with no compromise in functionality. Since the early years of Business Intelligence, there have been five major generations, or cycles, of computing: mainframes, miniframes, PCs, desktop Internet computing and mobile Internet computing (or simply mobile computing). Mobile computing, the fifth computing generation, has already had a huge impact and adoption rate, mobile devices are becoming the world's dominant computing platform. Business Intelligence (BI) is a natural fit for mobile devices. More organizations are considering mobile BI not just for executives, but for a variety of workers who need access to data when not at their seats. The technology and devices have evolved to make mobile BI cost-effective to deploy—better smartphones, new tablets and faster networks have all changed the game for mobile programs. Yet, many challenges remain around architecting, implementing and managing programs successfully.

Mobile technology makes it possible for decision makers to make immediate decisions. Users can shift through enormous volumes of data on their handheld devices and convert this data into actionable insight. Within moments, information is accessible without sitting down and finding a place to plug in a laptop. And rapid decision making is the key to accelerating the profitability of business. In today's fast-changing, competitive business environment, it is imperative to provide immediate answers to both internal and external customers. With Mobile Intelligence, decision makers now have the power to make these decisions immediately. In the mobile intelligence era, businesses that don't yet exist may evolve into industry leaders. Moderately valuable apps that run on desktops may become hugely successful apps when fully applied to the mobile Internet.

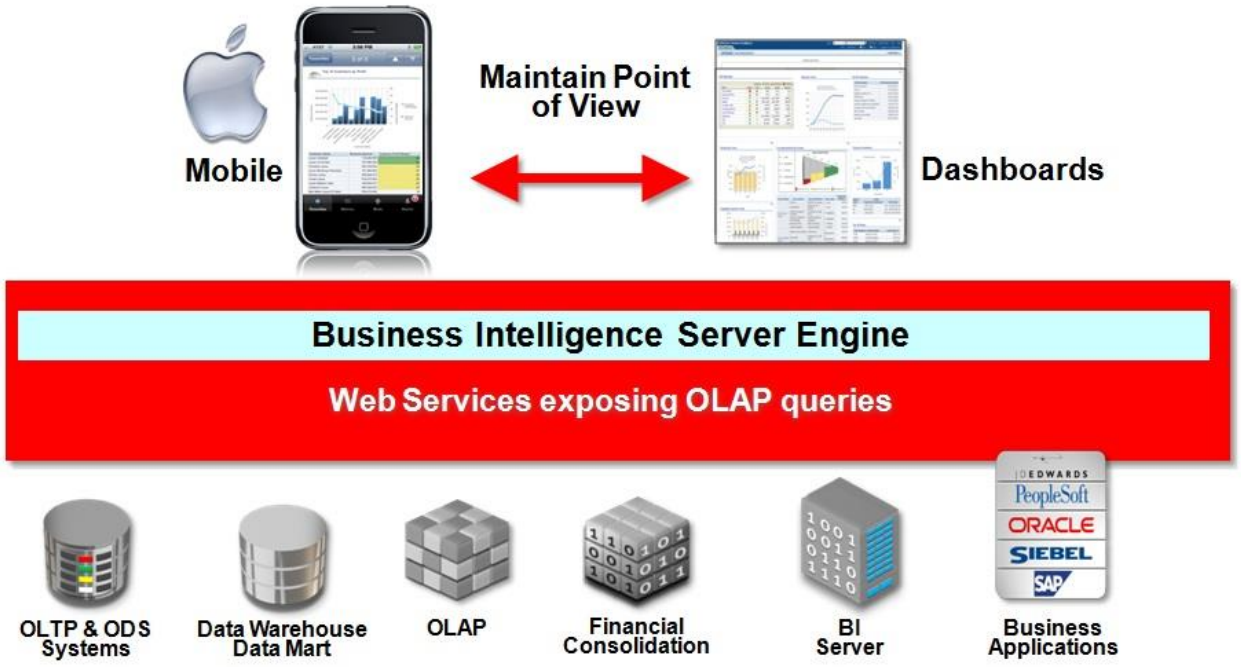
Organizations that stay with today's desktop-based information distribution models may become obsolete, outpaced by those organizations that choose to thrive on the mobile Internet. Organizations that embrace Mobile Intelligence will become leaner, faster, and be able to make smarter decisions resulting in more business, revenues, and competitive advantage.

Mobile intelligence is an abstract term much like Business Intelligence. Mobile intelligence is an encompassing term that expresses patterns that should be followed to deliver business intelligence tools on current mobile devices. With the dramatic growth of powerful smart phones and tablet devices, mobile intelligence is gaining dramatic traction in the business intelligence community. The mobile access to business data tends to become a rule rather than an exception Liu (2010). The idea of convergence between Business Intelligence and mobility is obvious within this context when more and more professionals start using smart phones and other mobile devices to keep up to date with business information.

Devices released in the last couple years are capable of storing large amounts of data, providing rich interactivity, accessing networks quickly, and visualizing complex data. This allows many best practices of business intelligence to be implemented on mobile devices. From our point of view, mobile intelligence can be simply defined as business intelligence for mobile devices. Business Intelligence patterns call for simple, interactive, provide real-time feedback, self-service (easy to learn) modules that

can be used by casual users. Mobile devices like the Windows Phone 7, Android based devices or the Apple iPhone are perfect examples of this Sajjad (2009), as we illustrated in Figure 2.

Figure 2. Example of Mobile Analytics framework.



Source: Vasiliev (2010).

One of the key differences between mobile devices (tablets, smart phones) and desktops is the large difference between input and interactivity. Mobile devices, for the most part, are not designed to be alternatives to quickly typing on a keyboard or the sensitivity a mouse provides for drawing. A large majority of the interactions on a mobile device are done to retrieve information.

Mobile devices like smartphones or tablets obviously provide virtual keyboards. Some even can accept external USB input devices like keyboards and mice. However, you will find that a very large majority of these devices are not used for data input because data input is much simpler and faster to do on a desktop. This is where the principles of Business Intelligence stand out in mobile intelligence implementations. You won't want to design an interface that requires frequent tapping, large text input, non-validated data, or non-real-time feedback as this will frustrate the casual user. For example, if you wrote an application that provided real-time analytics on a baseball game, you could have the user input the data play-by-play or implement a service that does this for them. This way the user focuses on only consuming business intelligence interactions.

Mobile intelligence does not have to be delivered as a "business intelligence only" application. Mobile Intelligence can be surfaced as modules that are part of an application or enhance additional functionality. For example, a Windows Phone 7 application could be created that manages web domains for an infrastructure engineer. The main features of the application would be the domain management functionality. The core functionality could be extended with predictive analytics to determine if the hardware can handle future visitor loads and possibly recommend a hardware upgrade. Those interactive applications on mobile devices are revolutionizing information dissemination and consumption. And the alignment of business data, analytics, and mobile computing is transforming business processes.

Mobile Intelligence is positioned to change how organizations deliver, consume, and act on information. Without 24/7 convenient access to business information, decisions and actions get postponed, causing bottlenecks and delays. These restrictions and delays are blown away with mobile intelligence, which allows heuristic analysis and decision making wherever a decision is required. As a segment of those professional contributing to the increase of need and demand for mobile Business Intelligence, the executive management stuff are also adapting especially through not being limited to their desks and office-based PCs. Together with the ubiquitous remote access, this reason makes the traditional executive workspace obsolete – now the whole world tends to becomes an office Kuntze (2010).

Beyond the access to many collections of data about corporate performance regardless of geographic location, time of day, and content delivery platform, the big advantages of the mobile tools for Business Intelligence are the business agility, the possibility to make dynamic and quality decisions based on real-time data and the promised standardization. In fact, as standardization has emerged and performance gaps have stabilized, the ease of use of mobile Business Intelligence applications has increased in parallel, reducing learning curves for the executive audience of mobile Business Intelligence devices Pedro (2004). For power-users, reports have become customizable on all types of mobile devices. Now Business Intelligence data can be consumed anywhere in the world on any wireless-enabled device including:

- Cell phones
- BlackBerrys
- iPhones
- Various types of PDAs
- Pocket PCs.

Based on these considerations we can reveal two concepts related to mobile intelligence framework's throughoutput:

- **Decision sweet spots**: these spots are locations, such as a commuter train, an aisle in a store, a line in a factory, or a retail floor. Business people need to be able to make data-driven decisions in the sweet spot, rather than delay due to a lack of information or analysis capabilities.
- **Decision windows of opportunity**: this window exists when a choice or action can be made to maximize an impact. The longer it takes someone to get to the information and completely evaluate the situation, the greater the chance of missing an opportunity. And delays risk the loss of a sale or customer.

The basic components of mobile intelligence tools for decision support systems include: *database*, *model base*, *method base*, *man-machine interface* and *intelligent components*. These basic components can be composed of two types of system structure: *four base structure* and the *fusion structure*.

This system is able to provide decision makers the necessary data, information and background materials to provide decision-making, to help clear goals and decision-making problem identification, to establish or modify the decision-making model, to provide a range of options, and a variety of program evaluation and selection excellent features of human-computer interaction analysis, comparison and to determine, to provide the necessary support for the correct and effective decision-making process.

Essential questions are finding a quick answer through mobile intelligence, fact which was unimaginable in the past with the older technology. Nowadays, key users can find answers to these important questions by simply securely accessing their smart-phones, PDAs, tablets, etc as we have illustrated in the Figure 2.

Figure 2: Mobile Intelligence tools provide answers to important questions

Source: Vasiliev (2010)

4. Conclusion

Mobile technology makes it possible for people to make immediate decisions. Key users can shift through enormous volumes of data on their handheld devices and convert this data into actionable insight. Within moments, information is accessible without sitting down and finding a place to plug in a laptop. And rapid decision making is the key to accelerating the profitability of business. In today's fast-changing, competitive business environment, it is imperative to provide immediate answers to both internal and external customers. With mobile intelligence, decision makers now have the power to make these decisions immediately. In the mobile intelligence era, businesses that don't yet exist may evolve into industry leaders. Moderately valuable apps that run on desktops may become hugely successful apps when fully applied to the mobile Internet. Mobile intelligence shows strong growth potential and a bright future in the field of mobile decision support systems applications.

The next generation services like Facebook, YouTube, or Twitter hasn't been invented yet, but it will be designed as mobile applications. Organizations that stay with today's desktop-based information distribution models may become obsolete, outpaced by those organizations that choose to thrive on the mobile Internet. Organizations that embrace mobile intelligence will become leaner, faster, and be able to make smarter decisions resulting in more business, revenues, and competitive advantage.

5. References

- Airinei, D., Homocianu, D. (2010) The Mobile Business Intelligence Challenge, *Economy Informatics*, vol. 10.
- Bâră, A., Lungu, I., Velicanu, M. , Diaconița, V., Botha., I. (2008) Extracting Data from Virtual Data Warehouses–a Practical Aproach of How To Improve Query Performance, *The 7th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases (AIKED'08), University of Cambridge, UK.*
- Bart Czernicki. (2010) *Silverlight 4 Business Intelligence Software*. Apress.
- Kuntze, N., Rieke, R., Diederich, G., Sethmann, R., Sohr, K., Mustafa, T., Detken, K.-O. (2010) Secure Mobile Business Information Processing, Embedded and Ubiquitous Computing (EUC). *2010 IEEE/IFIP 8th International Conference*.
- Gert H.N. Laursen, Jesper Thorlund. (2010) *Business Analytics for Managers: Taking Business Intelligence beyond Reporting*. John Wiley & Sons.

- Qiang Liu, Lixin Diao, Guangcan Tu. (2010) The Application of Artificial Intelligence in Mobile Learning, System Science, Engineering Design and Manufacturing Informatization. (ICSEM), 2010 International Conference.
- Pedro, J. S., Burstein, P., Cao, P., Churilov, L., Zaslavsky, A., Wassertheil, J. (2004) Mobile Decision Support for Triage in Emergency Departments, Decision Support in an Uncertain and Complex World. *The IFIP TC8/WG8.3 International Conference*. P. 714-723.
- Sajjad, B., Mir, A., Khawar, A. Bashir, F. and Tariq. A. (2009) An open source service oriented mobile business intelligence tool (MBIT), *Information and Communication Technologies. ICICT '09 International Conference*. Available at http://ieeexplore.ieee.org.
- Chou Tripuramallu. (2008) Research paper: *BI and ERP Integration*, Department of CIS. Available at http://ieeexplore.ieee.org.
- Yuli Vasiliev (2010) *Oracle Business Intelligence: The Condensed Guide to Analysis and Reporting*, Packt Publishing.

THE IMPORTANCE OF AUTOMATION IN THE DATA WAREHOUSING ENVIRONMENT – A CASE STUDY

NAGY Ilona Mariana
Ph.D. student, Faculty of Economics and Business Administration/Department of Economy Informatics, "Babes-Bolyai" University, Cluj-Napoca, Romania, mariana.nagy@econ.ubbcluj.ro

Abstract: *Data warehousing solution development represents a challenging task which requires the employment of considerable resources on behalf of enterprises and sustained commitment from the stakeholders. Costs derive mostly from the amount of time invested in the design and physical implementation of these large projects, time that may be decreased through the automation of certain processes. In this paper we analyse the importance of automating data structure creation and corresponding ETL processes, define key concepts in the data warehousing technology, and propose a framework for semi-automated implementation of the multidimensional models in the analytical environment from technical metadata.*

Key words: data warehouse; automatic data structure generation; technical metadata; multidimensional model

JEL classification: Y80

1. Introduction

Data warehousing solutions development represent complex endeavours that involve elaborated design and implementation activities, which lead in most cases to extended delivery times and increased costs. These types of projects also require the employment of varied resources within enterprises as well as sustained commitment from the stakeholders. Due to their well acknowledged complexity, given mostly by the necessary enterprise-wide data integration, some project implementation attempts are prone to failure. Failure in this case may involve situations in which the final costs are over the initial estimated and allocated budget, the delivery schedules are exceeded, or some project costs are unjustified (Adelman and Moss 2000). At the same time, companies that provide data warehousing solutions development and customization need to stay competitive on the market with regard to implementation, maintenance and other similar activities. This competitiveness is ensured significantly by small costs of development and short delivery times; therefore, reducing some of the involved costs augments the chances of successful solution delivery and increases business satisfaction.

Given these circumstances, the automation of certain processes in the data warehousing environment increases the chances of successful time- and cost-bound development. Thus, we introduce a framework for semi-automated implementation of data warehousing processes, with focus on the automated generation of data structures from technical metadata, and argue the importance of automation in the analytical systems.

The paper is structured as follows: in section 2 we present the fundamentals of the data warehousing technology and define several related concepts; in section 3 we present the generic architecture and enlist the data warehousing processes and discuss their appropriateness for automation; in section 4 we introduce a framework for automating the implementation of a multidimensional data model in the analytical environment. In the end we acknowledge related research work by presenting a brief literature review and outline our conclusions.

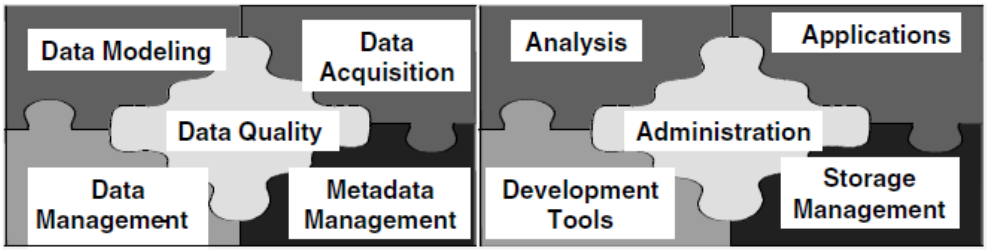
2. Data Warehousing: Fundamentals and Related Concepts

Data warehousing is defined as "a collection of decision support technologies, aimed at enabling the knowledge worker (executive, manager, and analyst) to make better and faster decisions" (Chaudhuri and Dayal 1997); as "a coordinated, architected, and periodic copying of data from various sources, both inside and outside the enterprise, into an environment optimized for analytical and informational processing" (Hammergren and Simon 2009); or as "the process of creating a repository of enterprise-wide data that is off loaded from production databases and is stored in a separate database that can be queried, reported and analysed" (van den Hoven 1998). All these definitions emphasize the fact that data warehousing represents rather a technology or a process that covers a wide range of activities involved in

the development of an analytical solution, than a single product. The activities involved in the creation of a data warehouse expand from visioning the idea of the final product, gathering valuable information, planning its evolution, and choosing the suitable development tools, to building and deploying the actual data warehouse, and employing maintenance activities (SDG Computing, Inc. 2007).

Data warehousing and the data warehouse concepts are presented thoroughly by Ponniah in (Ponniah 2001). He makes no distinction between the two concepts, and emphasizes the fact that the data warehouse represents "a user-centric computing environment where users can find structured information which can help them in the decision-making process; it is not a software or hardware product, which can be bought to provide strategic information". We argue that a clear differentiation should be made between the data warehousing technology and the data warehouse, as defined by W.H. Inmon (Inmon, 1995), also known as the "father of data warehousing": the data warehousing represents a technology that covers a wide range of activities and processes, while the data warehouse is a collection of data stored in a particular environment. However, we adhere to Ponniah's vision of data warehousing and acknowledge the statement according to which the data warehousing is a flexible and interactive, one hundred per cent user-driven environment for data analysis and decision support, which provides the ability to discover answers to complex and unpredictable questions, as well as a blend of various technologies, which go from data extraction, loading and transformation, to data storage, manipulation and presentation to the end-user, as depicted in Figure 12.

Figure 12. Data Warehousing: A blend of Technologies (Source: (Ponniah 2001))



Since the data warehousing represents a blend of technologies it has several related concepts, which we introduce below.

Data Warehouse

A data warehouse is defined in the context of the data warehousing technology as the area of storage where data is collected in order to be analysed; as a "collection of integrated, subject-oriented databases designed to support decision enhancing activities, where each unit of data is relevant to some moment in time" (McFadden and Watson 1996); or as a "copy of transaction data specifically structured for querying and analysis" (Kimball 1996). These similar definitions state clearly that the data warehouse integrates enterprise-wide data with the main goal of supporting the decision-making activity. W.H. Inmon (Inmon, 1995) gives a comprehensive definition of the data warehouse, as "a subject oriented, integrated, non-volatile, and time variant collection of data in support of management's decisions". These characteristics have been acknowledged universally in the academic environment and the industry and form the essence of its purpose (Inmon, 1996):

- the subject-oriented characteristic refers to the information being presented according to specific subjects or areas of interest, meaning that all data elements that relate to the same real-world event are linked together;
- the integration of data from various heterogeneous sources is one of the primary purposes of data warehousing as better decisions are made on complete information, gathered and integrated from enterprise-wide applications as well as other external sources;
- the non-volatility aspect refers to the stability of information in the data warehouse environment, which doesn't change every time an operational process is executed as it is not subject to record level update;
- the data warehouse is a time-variant environment, meaning that it stores historical data and not just current values of the transactions executed in the operational environment. This feature defines its purpose: integration of analysis and decision-making information over a period of

years. The time-variant nature of the data warehouse provides complex analysis of the past relating the information to the present and enabling forecast for the future.

Data warehouses are important for companies and their management in that they provide a foundation for decision making, by enabling an enterprise-wide view of data, in the first place. The collection of structured data which can be analysed and visualized in many different ways with the help of business intelligence tools forms the foundation for making better information-driven decisions. Thus, the enterprise-wide information is available in a structured and consistent form.

Data Mart

A data mart represents a specific, subject-oriented repository of data designed to answer specific questions. According to (Inmon, Imhoff and Sousa, 2001) a data mart is a "a subset of a data warehouse that has been customized to fit the needs of a department" that usually contains a small amount of detailed data and a larger portion of summarized data. In contrast, a data warehouse is seen as a single organizational repository of enterprise wide data across many or all subject areas. However, Kimball (Kimball, 1996) uses a series of individually developed data marts connected by a type of architecture known as "the bus-architecture" to provide the data warehousing functionally. Therefore, multiple data marts are used to serve the needs of multiple business units or departments from within an enterprise, such as sales, marketing, operations, collections, and accounting, etc.

OLAP (On-Line Analytical Processing)

According to (OLAP Council, 1997), "On-Line Analytical Processing is a category of software technology that enables analysts, managers and executives to gain insight into data through fast, consistent, interactive access in a wide variety of possible views of information that has been transformed from raw data to reflect the real dimensionality of the enterprise as understood by the user". The relationship between data warehouses and OLAP is complementary: while data warehouses are designed to store and manage data, OLAP transforms this data into strategic information by enabling a series of processing options which range from basic navigation and browsing (i.e. "slice and dice" operations), to complex calculations, and analysis of time series. OLAP uses a multidimensional view of aggregate data to provide quick access to strategic information for analysis purposes, whereas the data warehouse is based on relational technology.

The most fundamental characteristics of OLAP, as presented in (Ponniah 2001), are: enables users to have a multidimensional and logical view of the data in the data warehouse environment; facilitates interactive query and complex analysis for the users; allows users to perform drill-down and roll-up operations in order to analyze data at different granularity levels; provide the ability to perform complex calculations and present the results in a meaningful way (e.g. charts and graphs), etc.

Metadata

Metadata has been traditionally defined as "data about data". However, this definition is no longer valid in computer science and information systems especially, since its scope and usage has surpassed this limited view and interpretation. According to (Hay 2006), metadata is "the data that describes the structure of an organization's use of information and the system it uses to manage that information". Thus, metadata offers a broader view of the enterprise, defining how the data, activities and people are understood by it (how this data is represented in a structured manner, who can access the data and why, where the data is stored and the processes involved in its retrieval and manipulation, and the motivation and business rules which provide semantics to the data).

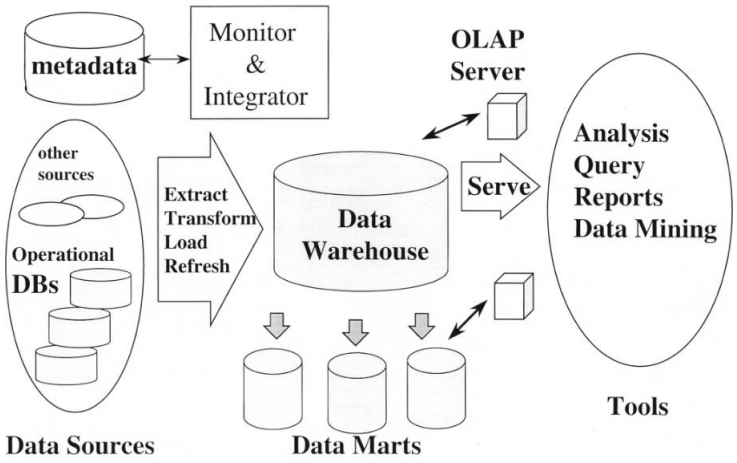
The enterprise's informational systems contain various types of metadata, from business and technical metadata, static and dynamic metadata, to descriptive, structural and administrative ones. The technical metadata, which we employ mostly in the automation processes of our research paper, is generated and used by the Information Technology systems, and it includes all the details on all the layers of the system, both at physical and logical levels; at the same time it is seen as structured information used by automated processes (Brand, Daly and Meyers 2003). According to Kimball (Kimball and Ross, 2002) "technical metadata defines the objects and processes in a system from a technical point of view, and includes data structures, databases, dimensions, measures and data mining models".

Metadata has an essential role in the analytical environment's data interpretation, by exposing a layer of semantic data – information- and meaning to the end user. In the data warehouse environment metadata helps facilitate the understanding of the enterprise's data assets, by solving the problem of early decision support systems – the information used and not the data supplied, needs to be addressed and encouraged. Its management helps minimizing the efforts of data warehouse administration and improves the extraction of information from the analytical environment.

3. The Data Warehousing Architecture and Processes

In order to get a broader understanding of how the proposed framework is integrated in the data warehousing system architecture, we use this section to present briefly the main architectural layers, components and processes (see Figure 13), along with some defining characteristics.

Figure 13. The Data Warehousing Environment Architecture (Source: (Segall 2004))



The main layers of the data warehousing environment include the data acquisition, the data warehouse and the data mart layers. In (Inmon, 1996), an additional operational data store layer is considered, designed to store data in a normalized form and to serve the operational reporting needs of the enterprise.

The data sources from which most of the data warehousing environment's data is extracted are represented by the operational systems of the enterprise. These often heterogeneous core systems that support day-to-day business operations (Inmon, Imhoff and Sousa, 2001) determine the success or failure of the data warehousing environment, as they are responsible for supplying the richness in data needed to understand the business and the history needed to judge the evolution of the business (Imhoff 1999). Operational source systems have certain particularities, which make them unsuitable for decision support and argues the necessity of a data warehousing environment within an enterprise (Imhoff 1999): most of the operational systems were not designed with data integration in mind, but built to perform a specific function; the linkage between different operational systems is weak (at best they pass a very limited amount of data between them); operational systems are not able to handle large amount of history, they are responsible for representing the current state of affairs. Besides these source systems, the data warehousing includes other external data sources, such as flat files, etc.

As discussed in the previous section, the data warehouse represents the area of storage where enterprise-wide data is integrated. The main roles of the data warehouse, defined in (Inmon, Imhoff and Sousa, 2001), are: the delivery of a common view of the enterprise data; supporting flexibility in how the data is further interpreted or used by consumers and providing a stable source of constant, consistent and reliable historical information; growing/expansion capabilities in order to store the enormous quantity of information generated by daily operational activities; serving of many information consumers, unlike the data mart designed to serve mostly top-management users.

The data mart layer portrays the dimensional model within the data warehousing environment. Data Marts are the primary structures used in analysis and reporting, they contain mostly aggregated data and enable the manipulation of this data with business logic (e.g. definition of key performance indicators, etc.). The Data Mart layer, in the Corporate Information Factory view, is implemented as an extension of the data warehouse and not as an alternative, designed to enable classical decision support, including data mining and data visualization processes (Imhoff 1999).

The Operational Data Store (ODS), as defined by Inmon in (Inmon, Imhoff and Battas, 1999), is "a dynamic architectural construct specifically designed for doing high-speed, integrated operational

processing". It can actually be seen as the bridge between the transactional and the analytical processing environments. The transactional characteristics of the ODS are given by its data structure and operational-level processing capabilities, while its analytical properties are defined by the high-quality cleansed and integrated highly granular data stored. These properties give the ODS a dual role within the data warehousing environment: a basis for mission-critical systems, with high response time and high availability, and support for tactical decision-making processes. The ODS supports transaction-level analysis, by storing data records in a non-summarized, query accessible and permanent form, without enabling the management of historical information.

According to (UC4 Software 2008), there are five major processes in data warehousing:

- extraction of data from source systems;
- data staging or transformation processes;
- data loading in target structures;
- creation of data marts;
- reporting processes.

The first three are refered to as ETL (Extraction-Transformation-Loading) processes, and form an essential part of data acquisition layer. The extraction processes are designed to pull data from the enterprise's operational source systems or external files, and place it into termporary storage structures that construct the data staging or persistent staging area. From here, the transformation processes stage the data (e.g. different data formats are reconciled). The loading processes are rensponsible for storing the transformed data in data warehouse specific structures.

Some of these processes may be automated due to their technical character (i.e. they do not require the input of data architects and modelers in the physical implementation phase). The most often automated processes in the data warehousing environment are the extraction, transformation and loading ones. Because data is generated continously on operational and transactional systems, processes that gather this generated data and haul it into the analytical environment need to run, depending on the informational needs of the enterprise, on a daily, weekly or monthly basis. Thus, onces the structure of the processes has been defined and physically implemented in the system, periodical batch jobs may be used to enable the automation of ETL. This ensures that new data is permanently integrated in the warehousing environment, and that this is done without the system developers' intervention.

However, we claim that other processes such as data structures creation (e.g operational data store, data warehouse and data mart layers) may also be partially automated. The benefits of automation in this case refer to faster development times and cost reductions. As discussed previously, the data warehouse represents the single organizational repository of enterprise-wide data that stores information oriented to satisfy decision-making request. It therefore has some particular features which makes its design and implementation processes different from the ones used in transactional systems.

We dedicate the next section of this paper to presenting a framework for the automation of data structure generation in the warehousing environment, present its components and discuss their interaction. We focus on the generation of the data mart layer specific data structures, i.e. the multidimensional data model and the description of the corresponding prototype implementation.

4. Case Study – A Framework for Process Automation

The design of a framework for semi-automated generation of a multidimensional data model and the implementation of the prototype represent a highly challenging endeavour. The proposed framework is intended to offer a foundation for data structures implementation in the analytical environment based on technical metadata (generated in the data warehousing system or derived from the transactional systems that provide the actual data). We assert that under certain conditions, as explained below, the automated data structure generation processes is feasible.

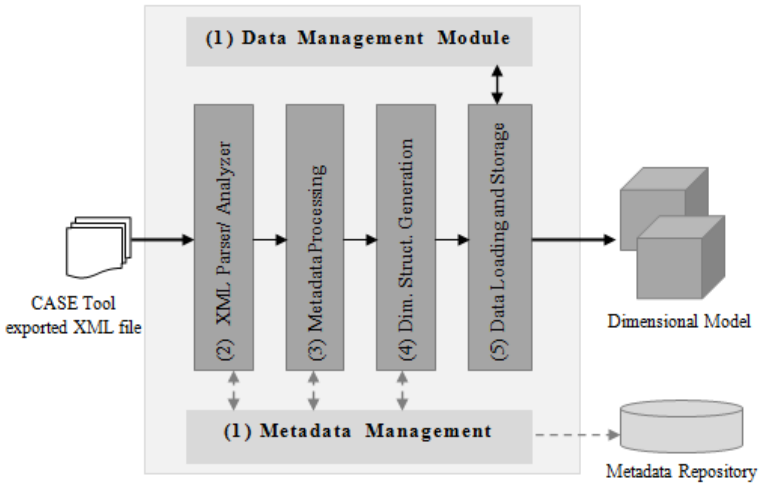
To get a broader understanding of how the framework components are designed with regard to the system architecture, we use this section to briefly clarify some important aspects:

- The framework for the semi-automated implementation of multidimensional data models is designed on the top storage layer of the enterprise data warehousing architecture (the data mart layer), also known as the Corporate Information Factory (CIF) (Inmon, et al. 2001).
- The data model which makes the case of the implementation must be provided in XML format as exported from a CASE (Computer Aided System Engineering) tool. Aspects of multidimensional data model representation as UML class diagram are presented in (Muntean 2007).

- In order the enable the automated generation of the data mart layer specific multidimensional data structures, several lower architectural level (i.e. data staging, operational data store and data warehouse) structures must already exists in the warehousing environment. The existence of these pre-requisite data structures is ensured by the extended version of the prototype, whose presentation does not make the case of this paper.

An overview of the proposed framework is presented in Figure 14.

Figure 14. Automation Framework Overview



The framework defines the basis for the automated prototype development that physically implements the multidimensional data model, and contains the following sub-components: (1) data and metadata management modules; (2) an XML parser/analyser component, responsible for reading, interpreting and mapping the class diagram to the relational permanent storage structures of the analytical environment; (3) a metadata processing component, which retrieves technical metadata from the repository; (4) a multidimensional data structure generation component, responsible for using the technical metadata and the logical mappings, and physically implement the data model in the relational analytical environment; and (5) data loading and permanent storage component, responsible for loading the data in the previously generated structures.

The metadata component contains information about the data warehouse, the corresponding data structures, the data stored in these structures, and the processes used for coordinating and managing all activities within the data warehousing environment. This component handles several types of metadata, presented in the previous sections, especially technical metadata in the context of the proposed framework, and has an essential role in the automation of the data structure generation processes. The other sub-components of the automation framework interact with the metadata management sub-component in order to retrieve the technical description of the data structures to be generated and to store the resulted metadata into the repository.

The automation of the data structure creation process is achieved with the help of the metadata management sub-component and the metadata repository. Extraction, transformation and loading processes are also generated from the technical metadata residing in the system, and may be schedule for periodical execution for populating the data warehouse. We mention however that the storage structures and corresponding ETL processes are generated in an initial phase, which means that they may be enhanced with more complex features (e.g. complex formulas, additional logical mappings, etc.) in a non-automated manner by data modellers. These enhancements require extended business knowledge along with capturing of business metadata into the repository, thus cannot be yet automatically incorporated in the resulting structures.

5. Literature review

The automation issues in the data warehousing environment have been previously addressed in the industry and academic world, most of them referring to ETL processes design, such as data extraction and population (Adzic, Fiore and Spelta 2001), data cleaning and integration (Jarke, et al. 1999) (Tziovara, Vassiliadis and Simitsis 2007), etc. There are several research works that focus on the

automation of conceptual or logical data warehouse design, such as (Phipps and Davis 2002) (Peralta, Illarze and Ruggia 2003), etc. Phipps and Davis propose an algorithm for automating the schema derivation from On-Line Transactional schema (i.e. the main source systems of the analytical data), named a "creation algorithm". They also present an "evaluation algorithm" following a user-driven requirements approach to select the candidate conceptual schemas with user queries. In (Peralta, Illarze and Ruggia 2003) the authors introduce a rule-based mechanism that automatically generates the relational data warehouse schema by applying existing design knowledge (a conceptual schema, non-functional requirements and mappings between the conceptual schema and the source database). In (Hahn, Sapia and Blaschka 2000) the authors present an automated generation tool for conceptual design model implementation from conceptual a graphical model, which comes closest to our idea. Our approach also considers aspects of automation of logical and physical mappings in the data warehousing environment, as treated in (Benkley, et al. 1995) (Castellanos, et al. 2009) (Rahm and Bernstein 2001).

To the best of our knowledge, automating the creation of data structures from technical metadata for the physical implementation of multidimensional data models has not yet been covered. We claim that given the existence technical metadata in the analytical system, along with external metadata files provided by data modellers (e.g. the UML class diagram of the star schema design), the proposed automation framework acts as a solid foundation for the development of a prototype that shortens considerably the manual repetitive work of data structures creation and enables the implementation of an initial multidimensional data models and corresponding ETL processes.

6. Conclusion

The framework presented in this paper is designed for automating certain processes of data warehousing data structures creation and corresponding ETL processes based on technical metadata provided by the underlying operational systems or external metadata files. The framework sub-components and their interaction act as a basis on which an automation tool may be created. This facilitates the implementation of the physical data model from existing and generated technical metadata and logical mappings of the underlying metadata repository. The main goal is to obtained faster implementation times in the data warehousing environment with reduced costs and delivery schedules. Our main contributions consist in the design of a framework which enables the interpretation of UML class diagrams exported in XML format and defines the ground for the implementation of the multidimensional model in the data warehousing environment from technical metadata.

7. Acknowledgements

Ph.D. scholarship, Project co-financed by the SECTORAL OPERATIONAL PROGRAM FOR HUMAN RESOURCES DEVELOPMENT 2007 – 2013;

Priority Axis 1. "Education and training in support for growth and development of a knowledge based society"; Key area of intervention 1.5: Doctoral and post-doctoral programs in support of research. Contract nr: POSDRU/88/1.5/S/60185 – "Innovative doctoral studies in a Knowledge Based Society" Babeş-Bolyai University, Cluj-Napoca, Romania

8. References

- Adelman, S., Moss, L. (2000, October). *Data Warehouse Failures.* Retrieved February 2012, from The Data Administration Newsletter: http://www.tdan.com
- Adzic, J., Fiore, V., Spelta, S. (2001). Data Warehouse Population Platform. *Proceedings of the VLDB 2001 International Workshop on Databases in Telecommunications II* . London, UK: Springer-Verlag.
- Ballard, C., Gomes, V., Hilz, G., Panthagani, M., Samuelsen, C. (2009). *Data Warehousing with the Informix Dynamic Server.* IBM Redbooks.
- Benkley, S., Fandozzi, J., Housman, E., Woodhouse, G. (1995). *Data element tool-based analysis (DELTA).* MITRE Technical Report.
- Brand, A., Daly, F., Meyers, B. (2003). *Metadata demystified.* (H. Pat, S. Parente, & K. Pirkey, Eds.) The Sheridan Press & NISO Press.
- Castellanos, M., Simitsis, A., Wilkinson, K., Dayal, U. (2009). Automating the loading of business process data warehouses. *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology* , (pp. 612-623).

- Chaudhuri, S., Dayal, U. (1997, May). An overview of data warehousing and OLAP technology. *ACM SIGMOD, 26*(1), p. 65-74.
- Colliat, G. (1996, September). OLAP, relational, and multidimensional database systems. *SIGMOD Rec., 25*(3), p. 64 - 69.
- Hahn, K., Sapia, C., Blaschka, M. (2000). Automatically Generating OLAP Schemata from Conceptual Graphical Models. *Proceedings of the ACM DOLAP 2000*, p. 9-16.
- Hammergren, T. C., Simon, A. R. (2009). *Data Warehousing for Dummies* (2nd Edition ed.). Indianapolis, Indiana, USA: Wiley Publishing, Inc.
- Hay, D. (2006). *Data Model Patterns - A Metadata Map.* San Francisco, California: Elsevier Inc.
- Imhoff, C. (1999, December). *The Corporate Information Factory.* Retrieved December 2011, from Information Management: http://www.information-management.com
- Inmon, W. (1995). "What is a Data Warehouse? *Prism Tech Topic, 1*(1).
- Inmon, W. (1996). *Building the Data Warehouse* (2nd Edition ed.). John Wiley & Sons, Inc.
- Inmon, W., Imhoff, C., Battas, G. (1999). *Building the Operational Data Store* (2 ed.). John Wiley and Sons, Inc.
- Inmon, W., Imhoff, C., Sousa, R. (2001). *Corporate Information Factory* (2 ed.). Wiley, Inc.
- Inmon, W., O'Neil, B., Fryman, L. (2008). *Business Metadata - Capturing Enterprise Knowledge.* Burlington, USA: Elsevier.
- Jarke, M., Quix, K., Blees, G., Lehmann, D., Michalk, G., Stierl, S. (1999). Improving OLTP data quality using data warehouse mechanisms. *Proceedings of the 1999 ACM SIGMOD international conference on Management of data.* Philadelphia, PA.
- Kimball, R. (1996). *The Data Warehouse Tookit: Practical Techniques for Building Dimensional Data Warehouses.* New-York: John Wiley and Sons, Inc.
- Kimball, R., Ross, M. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling.* New York City, U.S.A: John Wiley & Sons.
- Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., Becker, B. (2008). *The Data Warehouse Lifecycle Toolkit* (2nd ed.). New-York: Wiley, Inc.
- Knapp, D. (2009). *Practical Data Modeling with SAP NetWeaver® BW.* Bonn: SAP Press.
- McFadden, F., Watson, H. (1996). The world of data warehousing: issues and opportunities. *Journal of Data Warehousing, 1*(1), p. 61-71.
- Muntean, M. (2007). Implementation of the Multidimensional Modeling Concepts into Object-Relational Databases. *Revista Informatica Economică, 3*(43).
- OLAP Council. (1997). *What is OLAP?* White Paper .
- Peralta, V., Illarze, A., Ruggia, R. (2003). Towards the Automation of Data Warehouse Logical Design: a Rule-Based Approach. *In Proceedings of CAiSE Short Paper Proceedings.*
- Phipps, C., Davis, K. (2002). Automating Data Warehouse Conceptual Schema Design and Evaluation. *Proceedings of the 4th International Workshop on Design and Management of Data Warehouses (DMDW 2002).* Toronto, Canada.
- Ponniah, P. (2001). *Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals.* New-York: John Wiley & Sons, Inc.
- Rahm, E., Bernstein, P. (2001). A survey of approaches to automatic schema matching. *The VLDB Journal — The International Journal on Very Large Data Bases, 10*(4).
- SDG Computing, Inc. (2007, April). *Business Intelligence and Data Warehousing*. Retrieved October 2011, from http://www.sdgcomputing.com/index.htm
- Segall, R. (2004). Incorporating data mining and computer graphics for modeling of neural networks. *Kybernetes, 33*(8), p. 1258 - 1276.
- Staudt, M., Vaduva, A., Vetterli, T. (1999). *Metadata Management and Data Warehousing.* Zurich: University of Zurich.
- Tziovara, V., Vassiliadis, P., Simitsis, A. (2007). Deciding the Physical Implementation of ETL Workflows. *Proceedings of the ACM tenth international workshop on Data warehousing and OLAP*. New York, USA: ACM.
- UC4 Software. (2008). *Benefits of Automating Data Warehousing.* White Paper.
- van den Hoven, J. (1998). Data Warehousing: Bringing it All Together. *Information Systems Management, 15*(2), p. 1-5.

MODELING THE IMPACT OF ICT ON INNOVATION PERFORMANCE

PREDA Ana-Maria
Ph. D. Professor, Faculty of Management-Marketing, Romanian-American University, Bucharest, Romania
ampreda2007@yahoo.com

CRIŞAN Daniela Alexandra
Ph. D. Associate professor, Faculty of Computer Science for Business Management, Romanian-American University, Bucharest, Romania, dacrisan@yahoo.com

STĂNICĂ Justina Lavinia
Ph. D. Assistant professor, Faculty of Computer Science for Business Management, Romanian-American University, Bucharest, Romania, lavinia.stanica@rau.ro

Abstract: *The goal of this work is to model the impact of Information and Communication Technologies (ICT) on the Innovation performance for the European Union (EU) countries. In order to measure the overall innovation performance of each country we have used the Summary Innovation Index (SII), a composite indicator created on the initiative of the European Commission and computed yearly since 2001. The ICT performance was measured, at national level, using another composite indicator – the Networked Readiness Index (NRI) – assessed yearly since 2002 by the World Economic Forum, in collaboration with INSEAD. We have investigated two approaches: (1) a unifactorial linear regression between the two composite indexes: SII and NRI; (2) a unifactorial linear regression between SII and each of the four subindexes of NRI (Environment, Readiness, Usage, and Impact).*

Key words: ICT, innovation performance, NRI, SII

JEL classification: O 11

1. Introduction

In October 2010 the European Commission launched the Innovation Union initiative, one of the seven flagship initiatives of the Europe 2020 strategy for a "smart, sustainable and inclusive economy" (COM (2010)). The aim of this initiative is to support the development of research and innovation in Europe, by improving/changing the framework conditions and access to finance for these activities. The building of an Innovation Union is considered to be essential for the success of the Europe 2020 strategy as, according to specialists (European Commission, 2011), Europe is in a state of "innovation emergency".

On March, 2, 2012 the European Council stressed the importance of research and innovation as drivers of Europe's growth and jobs (European Commission, 2012). The Council recognized that "good progress has been made in launching and implementing 30 out of the 34 Innovation Union commitments" and that "Commission has put forward all six legislative proposals announced in the Innovation Union". Moreover, the newly proposed €80 billion investment program for research and innovation for 2014-2020 – Horizon 2020 – is expected to boost innovation and increase its efficiency and social-effectiveness, thus reducing/closing the "innovation divide" in Europe.

ICT is one of the areas where investments in research and development proved to be very efficient in terms of innovation and growth (COM (2009)). ICT is "by far the largest R&D-investing sector of the economy", and its pervasive impact, "its innovation performance and global dynamics, confirm the central role ICT plays in the world economy, the EU economy and the EU's economic recovery" (JRC EUR, 2010). More recent studies, such as the "Innovation Union Competitiveness Report", show that "the extent to which advanced ICT infrastructure is available to the business world .... drives innovation and influences the very way businesses innovate" (European Commission, 2011). The report highlights the role of ICT in the European research and innovation landscape, its contribution to achieve the Europe 2020 target of investing 3% of GDP in R&D, and to reduce the EU-US / EU-other international competitors gap (e.g. in competitiveness and business R&D intensity).

Thus, the link between ICT and innovation gets stronger and stronger, and increases in complexity, being more and more politically recognized. In this context, our study investigates the possibilities to model the impact of ICT on innovation for the EU-27 countries, in order to find significant

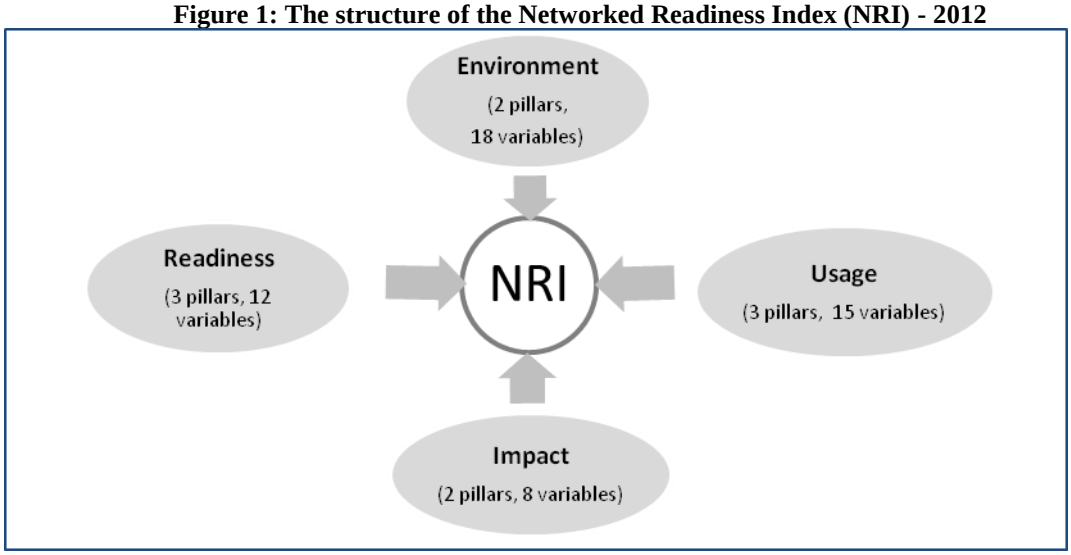
factors of influence for raising innovation. The first approach is to check the existence of a strong relationship between innovation and ICT, at country level. We'll use two composite indicators to evaluate ICT, respectively, the innovation performance of each country: Networked Readiness Index (NRI) and Summary Innovation Index (SII). If the result is positive, we'll extend our study by checking the existence of a relationship between innovation and some variables featuring ICT, at country level.

2. The Networked Readiness Index and the Summary Innovation Index - overview

The Networked Readiness Index (NRI) was introduced more than 10 years ago by the World Economic Forum, in collaboration with INSEAD, within the project "Global Information Technology Report" (GITR). The Report was published yearly since 2001, but the NRI framework remained stable since 2002. The aim of NRI is to measure how prepared countries are to access ICT and to use it effectively in order to enhance productivity, competitiveness, and well-being. NRI is a composite index based on dimensions, pillars, and variables. NRI-2012 is a new, evolved framework comprising both the main drivers of ICT readiness, and a new added dimension related to the ICT impacts on the economy and society (World Economic Forum, 2012). Thus, the new NRI gives a more comprehensive image on the efforts made by each country to provide access to ICT and to better use it, and on the effects of these efforts, respectively on the various aspects of the way ICT is transforming the economy and society. NRI has now four equal weighted dimensions (subindexes):

- the *"Environment"* for ICT – measures the general business and innovation environment, and the political and regulatory environment that facilitate the ICT access and use;
- the *"Readiness"* to use ICT – measures the ICT infrastructure level of development and the accessibility of digital content, the cost of accessing ICT (affordability), and the ability of a society (the skills) to effectively use ICT;
- the *"Usage"* of ICT – measures the extent and quality of ICT usage by all main stakeholders, respectively: individuals, businesses, and government;
- the *"Impacts"* of ICT – measures the economic and social effects of ICT (on competitiveness, and well-being) .

Figure 1: The structure of the Networked Readiness Index (NRI) - 2012



Source: made by authors

The subindexes Environment, Readiness and Usage are ICT "Drivers", while the Impact subindex reveals the effects ("outputs"/"results") of ICT on the economy and society. NRI-2012 includes 142 economies (3 more than NRI-2011), accounting for over 98 percent of the global GDP (World Economic Forum, 2012).

The Summary Innovation Index (SII) was created at the request of the European Council in Lisbon in 2000, and has been published annually, from 2001 to 2009, in the European Innovation Scoreboard (EIS). The aim of SII is to measure the overall innovation performance of an economy. After the launching of the Innovation Union initiative in 2010, the research and innovation performance of the EU Member States was monitored through the Innovation Union Scoreboard (IUS) that uses SII. Similar

to the NRI indicator, SII reflects both the drivers of innovation, and its effects/impacts on economy and society. SII is – like NRI – a composite index based on three categories (subindexes):

- the *"Enablers"* of innovation - capture the main drivers of innovation performance external to the firm (building blocks that allow innovation to take place);
- the *"Firm activities"* - cover the innovation efforts inside firms;
- the *"Outputs"* – refer to the effects of firms' innovation activities (the benefits of innovation).

Each category has pillars (named "dimensions"), and indicators (variables). SII structure comprises: 3 categories, 8 dimensions, and 25 indicators. SII indicator is calculated as the unweighted average of the re-scaled scores for all indicators.

Every year, the countries analyzed in the scoreboard are grouped into four classes according to their innovation performance measured by the SII indicator. The Innovation Union Scoreboard (IUS) 2011 (PRO INNO, 2012) uses 24 indicators, and analyses 34 European countries. It proposes the following classification for the EU-27 Member States:

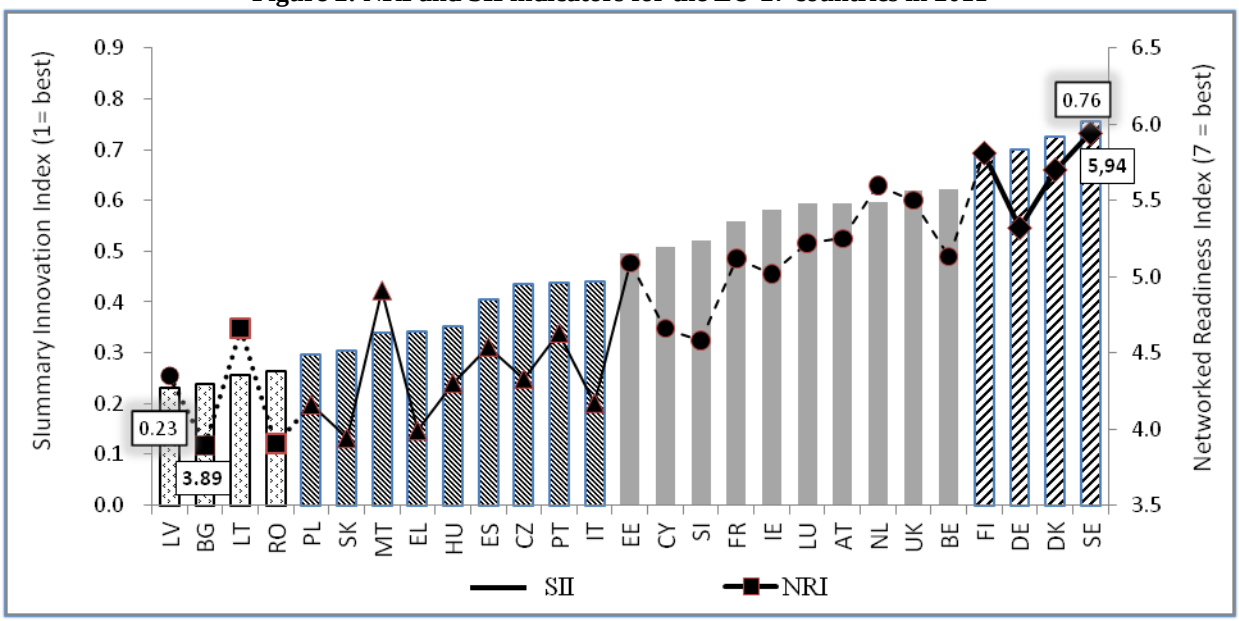
- *Innovation leaders* (Denmark, Finland, Germany and Sweden) – perform in innovation well above the EU-27 average;
- *Innovation followers* (Austria, Belgium, Cyprus, Estonia, France, Ireland, Luxembourg, Netherlands, Slovenia and the UK) – innovate below the leaders, but close to or above that of the EU-27;
- *Moderate innovators* (Czech Republic, Greece, Hungary, Italy, Malta, Poland, Portugal, Slovakia and Spain) – perform in innovation below the EU-27 average;
- *Modest innovators* (Bulgaria, Latvia, Lithuania and Romania) – innovate well below the EU-27 average (their innovation performance is less than half that of the EU-27).

3. Econometric tests

a. Modeling the relation Innovation – ICT for the 27 European Union countries

The first step in our analysis was to investigate the correlation of the two indicators – SII and NRI – for the EU-27 Member States. We have used data from the Innovation Union Scoreboard 2011 (published in February 2012) and the Global Information Technology Report 2012 (published in April 2012). Figure 2 shows SII and NRI values for each of the 27 EU countries, represented according to their innovation group ("Modest innovators" on the left side, and "Innovation leaders" on the right side).

Figure 2: NRI and SII indicators for the EU-27 countries in 2011



Source: Made by authors (IUS 2011, GITR 2012)

The graphics in Figure 2 reveals the gap between the best performers in terms of ICT/innovation (the "Nordic" countries: Sweden, Finland, and Denmark) and the worst performers (Romania, Bulgaria, Latvia, Lithuania, Poland, and Slovakia). NRI shows the depth of the digital divide within the EU-27

Member States (e.g. Sweden is 1.53 times more performing in ICT than Bulgaria). At the same time, SII shows the "innovation divide", the big differences in innovation performance among EU-27 economies (e.g. Sweden is 3.28 times more performing in innovation than Latvia, and 3.16 times more performing than Bulgaria).

We have tested the following unifactorial regression model:

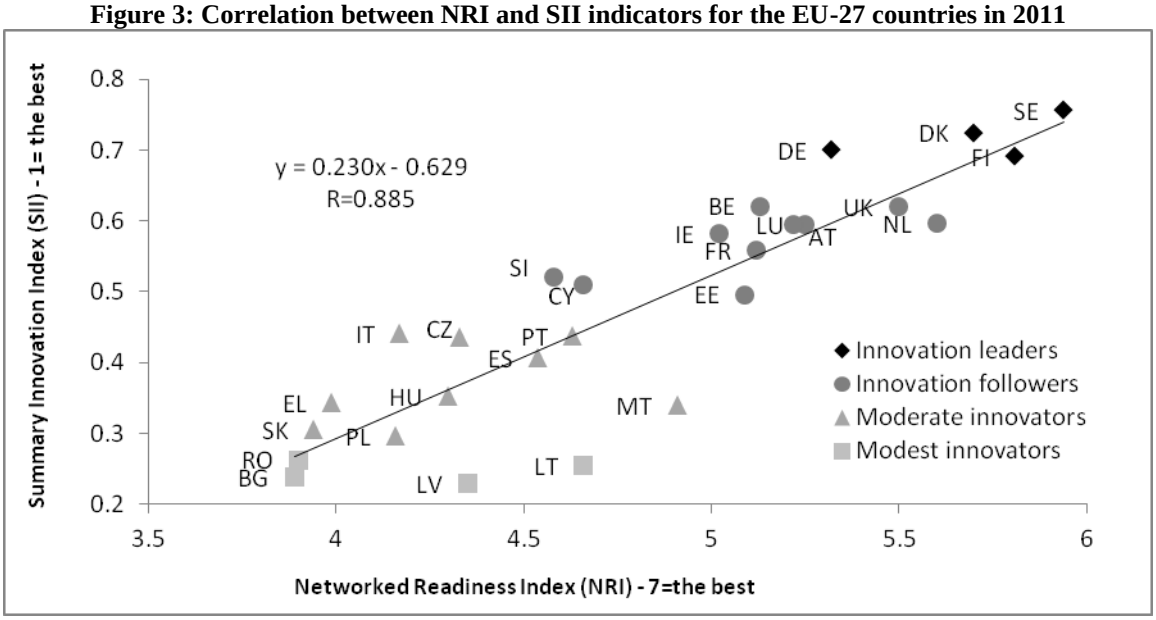
*SII = <slope> * NRI + <intercept>*

The statistical results are presented in detail in Annex. Table1. The regression model found was:

*SII = 0.230*NRI - 0.629*

Figure 3 proves the strong correlation between the two indicators, sustained by the high value of the coefficient of correlation: R= 0.885.

Figure 3: Correlation between NRI and SII indicators for the EU-27 countries in 2011



Source: Authors' calculations

In Figure 3 the EU-27 countries are represented according to their innovation group. It can be observed the higher position, in terms of both innovation and ICT, of the four "Innovation leaders" (on the top-right corner). The group of "Innovation followers" (with the exception of Slovenia and Cyprus) is rather homogenous and close to the EU trend. In this group, Netherlands has the best ICT performance, and Belgium the best innovation performance. In the "Moderate innovators" group, Italy is the leader in innovation (above the EU trend), while Malta is the leader in ICT (much below the EU trend). The "Modest innovators" are grouped at the bottom-left corner. Romania is the most innovative in its group, but far behind Lithuania and Latvia in ICT terms. Nevertheless, Romania is the closest "Modest innovator" to the EU trend. The regression model shows also that a 10-point increase in NRI induces an increase of the SII by approximately 1.7 points.

The second step in our analysis was to check the validity in time of the unifactorial regression model that depicts the correlation between SII and NRI. We have used a transversal analysis for the period 2007-2011. The results are presented in Table 1.

Table 1: Transversal analysis of the unifactorial linear regression between the global indicators SII and NRI at national level, for EU-27 countries, for the period 2007-2011

Year	**2007**	**2008**	**2009**	**2010**	**2011**
Slope	0.250	0.245	0.250	0.272	0.230
Intercept	-0.729	-0.702	-0.676	-0.763	-0.629
R	0.896	0.893	0.928	0.928	0.885

Source: Authors' calculations

As it can be noticed, there is a strong relationship between SII and NRI (coefficients of correlation greater than 0.88) in each of the 5 years of analysis. The parameters of the regression equation

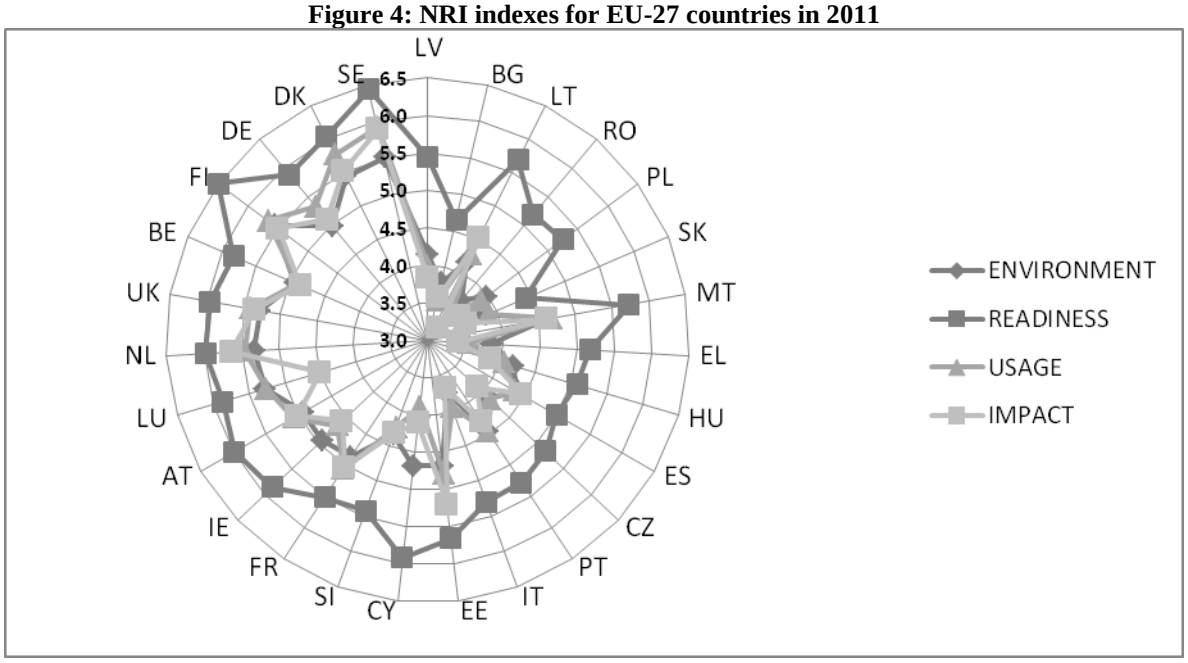
have changed in time, but insignificantly. The correlation between SII and NRI is confirmed by the data in each of the years from 2007 till now.

b. Modeling the relation Innovation – ICT (at structural level) for the 27 European Union countries in 2011

The third step in our analysis was to search in-depth – at structural level – the relationship between SII and NRI that we've proved in section (a). We've used the 4 subindexes of NRI (Environment, Readiness, Usage and Impact) and we've analyzed the relationship between SII and these 4 variables.

First, we've analyzed the values of the four NRI subindexes, in 2011, for all EU-27 countries.

Figure 4: NRI indexes for EU-27 countries in 2011



Source: Made by authors (GITR 2012)

The "Nordic countries" (SE, DK, FI – cluster 1: Innovation leaders) have, by far, the best performances: Sweden is the leader for Usage and Impact, and Finland is the leader for Environment and Readiness; Sweden and Denmark are in the "top 3 performers" for all NRI indexes; Sweden, Finland and Denmark are in the "top 3 performers" for 3 NRI indexes. At the opposite side, countries of cluster 4 – Modest innovators – have very low performances: Romania is the worst performer for 3 NRI indexes (Environment, Usage, and Impact), and Bulgaria is in "top 3 worst performers" for 3 NRI indexes (Environment, Readiness, and Usage). Table 2 shows, in more detail, the dispersion of values for EU-27 countries, for each of the NRI indexes, in 2011.

Table 2: Best and worst performers for each of the four NRI subindexes in 2011

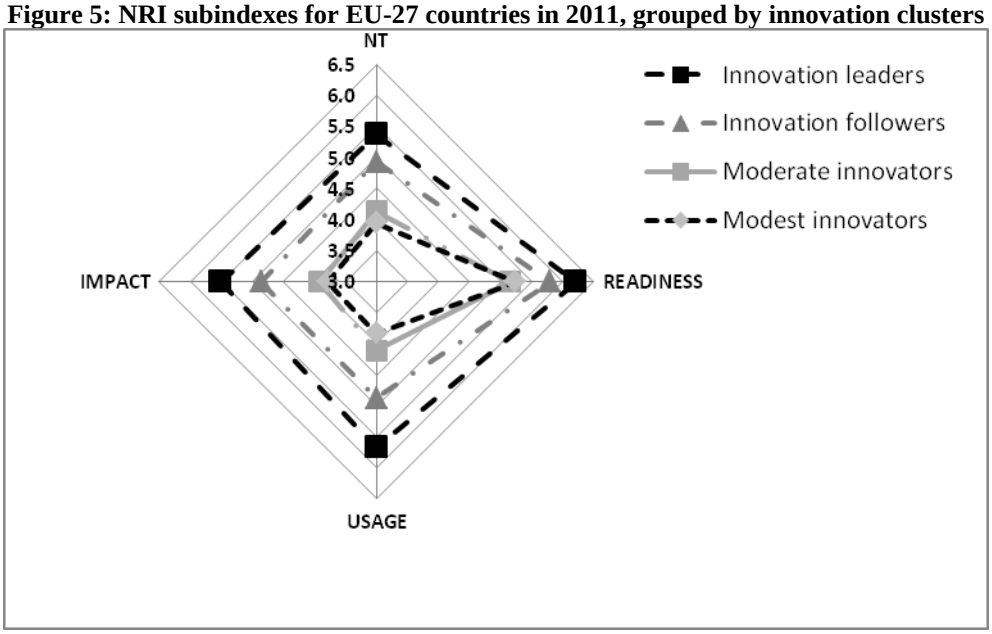
NRI index	Best performers			Worst performers			Comments
Environment	FI cluster 1	SE cluster 1	DK cluster 1	RO cluster 4	IT cluster 3	BG cluster 4	- 6 countries have values above 5 (cluster 1 and 2) - 15 countries have values between 4 and 5 - 6 countries have values below 4 (cluster 3 and 4)
	max=5.56			min=3.69			
Readiness	FI cluster 1	SE cluster 1	DK cluster 1	SK cluster 3	BG cluster 4	ES cluster 3	Relative homogenous values: - 3 countries have values above 6 (cluster 1) - 21 countries have values between 5 and 6 - 3 countries have values below 5 (cluster 4)
	max=6.5			min=4.43			
	SE	DK	FI	RO	BG	EL	Rather equally distributed

Usage	cluster 1 max=5.92	cluster 1	cluster 1	cluster 4 min=3.5	cluster 4	cluster 3	countries on the intervals: - above 5 (9 countries: clusters 1, 2) - between 4 and 5 (10 countries: clusters 2,3) - below 4 (8 countries: clusters 3,4)
Impact	SE cluster 1 max=6.5	NL cluster 2	DK cluster 1	RO cluster 4 min=3.21	EL cluster 3	PL cluster 3	Equally distributed number of countries on the intervals: - above 5 (9 countries) - between 4 and 5 (9 countries) - below 4 (9 countries)

Source: Made by authors
Legend: cluster 1 = Innovation leaders; cluster 2 = Innovation followers; cluster 3 = Moderate innovators; cluster 4 = Modest innovators

In Figure 5 we've represented the values of NRI subindexes for each of the 4 innovation clusters. We've used the average score of the NRI subindexes of the countries within each innovation cluster.

Figure 5: NRI subindexes for EU-27 countries in 2011, grouped by innovation clusters



Source: Made by authors (IUS 2011, GITR 2012)

Table 3 presents a synthesis of the analysis made on the NRI subindexes values for each innovation cluster. The table highlights the strengths and weaknesses in each innovation cluster, and which of the NRI subindexes could be improved for certain countries.

Table 3: NRI subindexes, in 2011, for EU-27 countries, analyzed by innovation cluster

Innovation cluster	Comments
Innovation leaders (4 countries)	• Belong to the "best 3 (4) performers" for all NRI indexes • Best results for indexes: Usage and Impact • The worst result = DE (for Environment) • The most important factors for DE are: Environment and Readiness
Innovation followers (10 countries)	• NL in "best 3 performers" for Impact index • More than 50% of countries have values in the highest range (over 5) for the indexes: Impact, Environment, and Usage • The worst result = IT (for Usage)
Moderate innovators (9 countries)	• At least one country belongs to the "worst 3(4) performers" for all NRI indexes: - SK, ES – for Readiness; - EL – for Usage; - IT – for Environment;

	- EL, PL, MT – for Impact • The worst results: - Impact index: 6 countries have values in the lowest range (below 4)) - Usage, and Environment indexes: 4 countries have values in the lowest range - (below 4)
Modest innovators (4 countries)	• At least one country belongs to the "worst 3 performers" for all NRI indexes: - BG – for Readiness; - RO, BG – for Usage, and Environment; - RO – for Impact • The worst result: for Impact index (3 countries have values in the lowest range) • Critical factors: - for RO: Impact, Usage, Environment - for BG: Usage, Environment, Readiness

Source: Made by authors

Then, we've tested the relationship between SII and each of the NRI subindexes using a unifactorial regression model:

SII = <slope> * NRI subindex + <intercept>

The results are synthesized in Table 4.

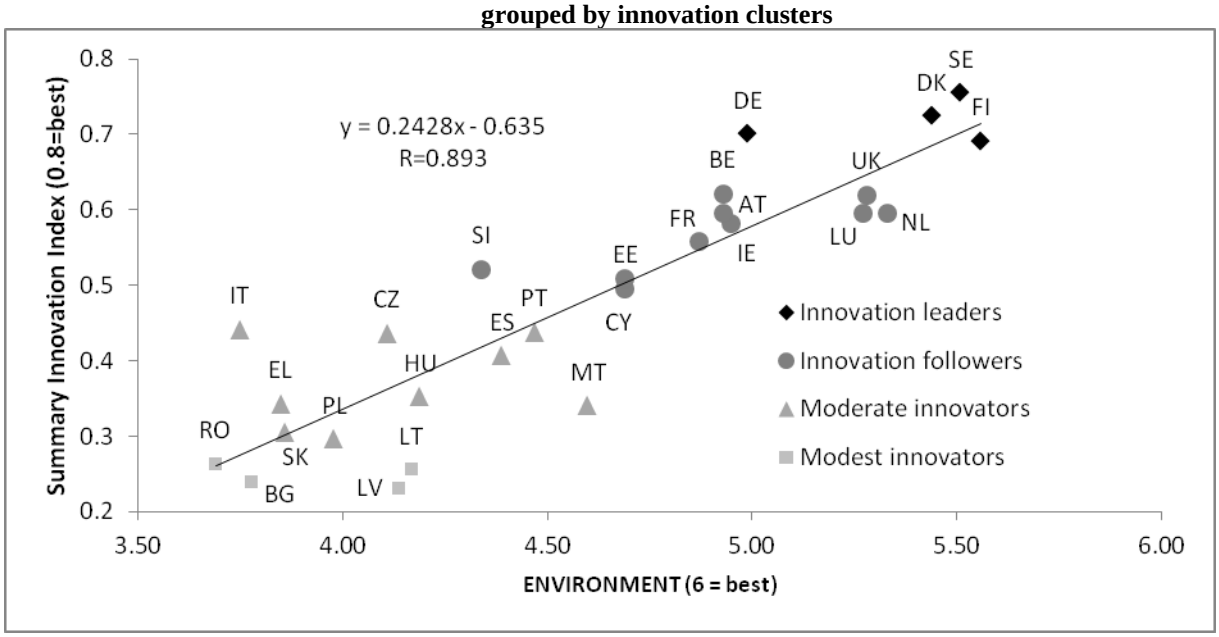
Table 4: The correlations between SII and NRI subindexes, at national level, for EU-27 countries, in 2011

NRI subindex	Environment	Readiness	Usage	Impact
Slope	0.243	0.255	0.195	0.023
Intercept	-0.635	-0.940	-0.416	0.106
R	0.893	0.781	0.880	0.834

Source: Authors 'calculations

We observe that there is a significant relationship between SII and each of the NRI subindexes (strong correlation proved by the correlation coefficients which have values above 0.78). The most powerful relationship is yield by the Environment factor (see fig.6).

Figure 6: The correlation between SII and the Environment subindex of NRI, for EU-27 countries, in 2011, grouped by innovation clusters



Source: Authors' calculations

Innovation leaders (excepting Finland) and Innovation followers (excepting United Kingdom, Netherlands, and Luxembourg) are close to the EU trend, with an innovation performance usually above the "expected value". Conversely, the Moderate innovators (excepting Czech Republic, Italy, and Spain), and the Modest innovators have an innovation performance well below that "expected". The graphics in

fig. 6 also shows that countries where the "*Environment for ICT*" (i.e. (1) the political and regulatory environment: legal system, intellectual property protection, software piracy, bureaucracy, and (2) the business and innovation environment: availability of latest technologies and of venture capital, fiscal system, easiness to start a business, enrollment in tertiary education, quality of management schools) is well developed have also a good innovation performance. This means that countries in the Modest innovators cluster (like Romania, Bulgaria) should make efforts to increase their "Environment for ICT" performance as a way to improve the overall innovation performance.

Similar conclusions can be formulated for the relationships between SII and the other 3 NRI subindexes:

The "*Usage of ICT*" (i.e. (1) individual usage: mobile phone/fixed broadband/mobile broadband subscriptions, internet usage by individuals/households, virtual social network, (2) business usage: firm-level technology absorption, capacity for innovation, PCT patents and applications, extent of business Internet use, extent of staff training, and (3) government usage: the number of on-line services provided by the government, government prioritization of ICT, and the importance given to ICT in its vision of the future) is a very influential factor to increase the innovation performance.

Also, the last ICT driver – "*Readiness to use ICT*" (i.e. (1) the ICT infrastructure and the accessibility of digital content: electricity production, mobile network coverage, international internet bandwidth, secure internet servers, (2) the affordability of ICT: mobile cellular tariffs, fixed broadband internet tariffs, internet and telephony sector competition, and (3) the skills for ICT: quality of educational system/math and science education, secondary education enrollment, adult literacy) – has a significant positive influence on the innovation performance of the EU countries.

Last, but not least, the "*Impact of ICT*" (i.e. (1) economic impacts: impact on new services and products, ICT PCT patents, and applications, impact on new organizational models, knowledge-intensive jobs, and (2) social impacts: impact on access to basic services, Internet access in schools, ICT use & government efficiency, e-participation index) could help in reducing the innovation divide between the "Innovation leaders" and the "Modest innovators".

4. Conclusions

As it was revealed in the Innovation Union Scoreboard 2011, EU still lags behind international leaders in innovation – US, Japan, South Korea – and has not closed its innovation gap. Also, there is a significant innovation gap (innovation divide) within EU Member States, especially between the "Innovation leaders" and "Modest Innovators". All these facts prove the need for an accelerated increase in the innovation performance of all EU countries, and mainly of those lagged behind. Finding the best solutions to increase innovation performance and to improve the national research and innovation system is now a "must" for all policy makers.

Our study revealed that ICT is closely related to the innovation performance, the relationship being modeled by a unifactorial linear regression. We've proved that the relationship is valid for all EU-27 countries, and for a long period of time (2007-2011). Moreover, we've analyzed the relationship between the innovation performance of EU-27 countries in 2011 and the "drivers of ICT" (environment, readiness, usage), respectively the "impacts of ICT". We've revealed the existence of a strong correlation between SII and each of the NRI subindexes, modeled through a unifactorial linear regression. These models showed that all ICT drivers and the ICT impact have a positive influence on the innovation performance, and could be considered as significant leverages for pushing up innovation. Improving all the aspects (indicators) included in the NRI subindexes could be a successful way to accelerate the growth of innovation in each EU country.

5. References

- COM (2009)116, *"A Strategy for ICT R&D and Innovation in Europe: Raising the Game"*. Available at http://ec.europa.eu/information_society/tl/research/documents/ict-rdi-strategy.pdf
- COM (2010) 546 final, *"Europe 2020 Flagship Initiative Innovation Union"*. Available at http://ec.europa.eu/research/innovation-union/pdf/innovation-union-communication_en.pdf
- European Commission (2010), "*Europe 2020 – A European strategy for smart, sustainable and inclusive growth"*. Available at http://www.i4cense.org/sites/default/files/Europe_2020.pdf
- European Commission (2011), *"Innovation Union Competitiveness Report - 2011 edition"*. Available at http://ec.europa.eu/research/innovation-union/index_en.cfm?section=competitiveness-report&year=2011

- European Commission (2012), "*European Council confirms research and innovation as drivers of growth and jobs".* Available at http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/12/153&format=PDF&aged=0&language=EN&guiLanguage=en
- JRC EUR 23832 EN (2009), "*The 2009 Report on R&D in ICT in the European Union".* Available at http://ftp.jrc.es/EURdoc/JRC49951.pdf
- JRC EUR 24320 EN (2010), "*The 2010 Report on R&D in ICT in the European Union".* Available at http://ftp.jrc.es/EURdoc/JRC57808.pdf.
- PRO INNO Europe® (2008), *"European innovation scoreboard (EIS) 2007",* PRO INNO Europe® Paper No. 6
- PRO INNO Europe® (2009), *"European innovation scoreboard (EIS) 2008: Comparative analysis of innovation performance",* PRO INNO Europe® Paper No. 10
- PRO INNO Europe® (2010), *"European Innovation Scoreboard (EIS) 2009: Comparative analysis of innovation performance*", PRO INNO Europe® Paper No. 15.
- PRO INNO Europe® (2011), "*Innovation Union Scoreboard (IUS) 2010: The Innovation Union's performance scoreboard for Research and Innovation".* Available at http://www.proinno-europe.eu/inno-metrics/page/innovation-union-scoreboard-2010
- PRO INNO Europe® (2012), *"Innovation Union Scoreboard (IUS) 2011: Research and Innovation Union Scoreboard".* Available at http://www.proinno-europe.eu/inno-metrics/page/innovation-union-scoreboard-2011
- World Economic Forum (2008), *"The Global Information Technology Report 2007-2008",* Available at http://networkedreadiness.com/gitr/main/previous/
- World Economic Forum (2009), "*The Global Information Technology Report 2008-2009".* Available at http://networkedreadiness.com/gitr/main/previous/
- World Economic Forum (2010), *"The Global Information Technology Report 2009-2010. ICT for Sustainability"*. Soumitra Dutta, Irene Mia, Ed. Oxford University Press, New York
- World Economic Forum (2011), *"The Global Information Technology Report 2010-2011. Transformations 2.0"*. Soumitra Dutta, Irene Mia, Ed. Oxford University Press, New York
- World Economic Forum (2012), "*The Global Information Technology Report (GITR) 2012 - Living in a Hyperconnected World*", Soumitra Dutta, Beñat Bilbao-Osorio, Ed. Oxford University Press, New York

Annex

Table 1. Statistical analysis of the unifactorial linear regression between global indicators SII and NRI, at national level, for EU-27 countries, in 2011

Regression Statistics	
Multiple R	0.885
R Square	0.784
Adjusted R Square	0.775
Standard Error	0.076
Observations	27

ANOVA

	df	SS	MS	F	Significance F
Regression	1	0.529	0.529	90.518	8.68E-10
Residual	25	0.146	0.006		
Total	26	0.675			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	-0.629	0.117	-5.362	1.46E-05	-0.870	-0.387
NRI	0.230	0.024	9.514	8.68E-10	0.181	0.280

*SII = 0.2304 * NRI - 0.6288*
R^2 = 0.784, strong correlation

Anova Analysis: Two statistical tests were used in order to verify the significance of the linear regression:
1. the Fisher-Snedecor test was used in order to determine whether the dependence between the selections of data is not random. For a significance level of $\alpha = 0.05$, one independent variable (NRI) and 27 observations (countries), we have 1 and 27-1-1=25 the two degrees of freedom, so the critical value of F is: $F_{critical}(0.05;1.25) = 4.241$. Since F-statistics = 90.518 >> $F_{critical}(0.05;1.25) = 4.241$, we conclude, at a significance level $\alpha = 0.05$, that the observed strong correlation between the variables did not occur by chance.
2. the "t" - Student test was used in order to determine whether the NRI parameter in the regression line is useful in estimating the SII value. Considering $\alpha = 0.05$ as before, the critical value of T distribution is: $T_{critical}(0.05;25) = 2.059$; since the T-statistics for the NRI variable is T-statistics(NRI)= 9.514> $T_{critical}(0.05;25) = 2.059$, we conclude that the estimation corresponding to the NRI variable is useful in the prediction line.

MOBILE DATA MANAGEMENT FOR BUSINESS INTELLIGENCE IN DISTRIBUTED ENVIRONMENT

PUICAN Florina Camelia
Ph. D. Student, Academy of Economic Studies / Department of Economic Informatics, Bucharest, Romania, puicanflorina@yahoo.com

MUNTEANU Adrian
Ph. D. Student, Academy of Economic Studies / Department of Economic Informatics, Bucharest, Romania, adrianm21@yahoo.com

ULARU Elena-Geanina
Ph. D. Student, Academy of Economic Studies / Department of Economic Informatics, Bucharest, Romania, ularugeanina@yahoo.com

Abstract: *The mobile systems' paradigm emerged because life cannot be imagined without their use both in personal and business purposes. The benefits of mobile technologies imply managing key information in real time, from mobile devices, the possibility to configure it in the easiest way and to use it for strategic decision with significant impact on business development. The aim of this paper is to describe the architecture of mobile data management software and the advantages obtained by managing a heterogeneous portfolio of mobile devices and operating systems for business intelligent applications.*

Keywords: advanced database systems, distributed systems, mobile data management, business intelligence

JEL classification: C88, D85, M15, M2

1. Introduction

The latest mobile technologies allow to display the applications and the collaboration between distributed environments. The real-time communication and transfer of information are the key factors in effective coordination and collaboration between dispersed environments. The development of mobile and wireless networks gained a substantial market share once with the awareness of their added value. Both the benefits brought up to society and revenue potential for businesses are enormous.. According to Gartner market share analysis (2011) and the estimates of the International Telecommunication Union, at the end of 2011, there were 6 billion mobile subscriptions. The global penetration reaches 87%, and 79% in the developing world. The mobile-broadband subscriptions grown 45% annually over the last four years and today there are twice as many mobile-broadband as fixedbroadband subscriptions.

Today, a topic of common interest is represented by the graphical view in terms of mobile applications. In business, mobile technologies are used by organizations in order to foster collaboration between their different activities. Private companies have a clear definition of their requirements: fast access to information, user-friendly data design, in an easy manner to use and the real time progress monitoring. Data visualization involves 3D, 4D technologies to nD modeling type, but the focus still remains on known technologies such as 2D type. If a business entity wants to have its applications to interact with applications from different business partners, new interfaces are needed to be built."This concept was defined as a *dynamic interaction*, meaning that any application can dynamically choose whom to interact with based on its best business interests". (Xu Yang, 2010)

Among the benefits of implementing mobile technologies, there are also communication networks and connections of different kind of softwares and platforms, but also the process of creating graphical applications for integrating in a collaborative environment. These mobile technologies have the following main goals: providing the communication platforms and communicating in real time, in terms of data transfer.

Devices that are suitable for mobile technologies are tablets - PC, mobile technologies 2D, 3D - to access visual data, 4D for progress monitoring, integration of Bluetooth and WLAN technology. Pocket devices - Pocket PC incorporates the 3G communications for real-time decisions, big screen

viewing information compared with PDAs, capturing videos and photos and video conferencing in real time.

Smartphones enable users to have access to charts and to navigate predefined drill paths that ensure additional details. There are mobile business intelligence providers that enable smartphone users to modify existing views (sort, compute, visualize...) as well as to manipulate information by updating remote applications or databases.

On the opposite side, tablet computers offer mobile business intelligence designers the possibility to provide a full range of oriented BI functionality. On a tablet computer, users can manipulate data in columns, explore data in any direction (versus being limited to predefined drill paths) as well as act on information via a variety of mechanisms, most notably email and annotations.

WLAN communication networks make the data transfer possible, with enhancements for 3D communication. Thus, information can be copied as desired, limiting the need of used memory device. WiMAX Networks (Word Wide Interoperability for Microwave Access) provide a more extended mobility than that of WLAN, but its costs are higher.

There are mobile services infrastructures which can provide a generic view for mobile users to lookup, access and execute web services over wireless broadcast networks. One of the most important issues of interest in mobile services field was highlighted by Athman Bouguettaya and it regards the "*access time* - the average time that has elapsed from the moment a client requests information up to the point when the required information is downloaded by the client". Another aspect concerns "*tuning time* - the amount of time that has been spent by a client listening to broadcast channels. This is used to determine the power consumed by the client to retrieve information" (Xu Yang, 2010). Nevertheless, in recent years, access time and tuning time were significantly diminished in clients favour.

"Mobile computing is the discipline that creates an information management platform, which is free from spatial and temporal constraints. The freedom from these constraints allows its users to access and process desired information from anywhere in the space. The state of the user, either static or mobile, does not affect the information management capability of the mobile platform". (Kumar, 2012)

An important aspect of mobile applications is the efficient data management. Although modern operating systems for mobile systems provide some support for databases, allowing applications to store permanent records in FLASH ROM, often more advanced techniques are needed. These include standard interfaces such as JDBC or ODBC, declarative query and manipulation languages, support for indexing and fast access to database replication. Therefore, database management systems offer smaller versions of them to mobile systems. For example, there are Oracle 9i and IBM DB2 Everyplace lite. These products offer low fixed functionality that can not be extended or customized. Most of the time, features offered by them are not needed for specific applications and there are situations when other unimplemented features may be essential.

Special requirements for specific applications of mobile systems exist at different levels in the database management systems. For example, some applications require query processing, while others require encryption of data stored, all of which can be made possible by a database management system. Other examples could be support for transactions, synchronization with a central database and support for special data types.

A single database management system can not meet all the above mentioned requests for mobile devices with limited resources. A viable approach could be an adapted database management system to allow the choice or combination of different modules from a set of components available on the device.

Improvements in data processing have led to the development of infrastructures for mobile applications using distributed environment. The research for the paper on distributed environment/distributed databases is aimed to create the basis of the proposal for mobile data management software architecture.

2. Mobile databases in distributed environment

Frameworks for distributed database systems and mobile ad-hoc networks (MANET) are aimed for mobile users, in order for them to have easy access to dispersed data available on mobile devices. A distributed system has directory nodes (DN) that store the database scheme of mobile devices andthe DN serves to transmit, if to a submissive user request can be answered partially or completely from the participating mobile database nodes (DBN). (Imielinski, 2003)

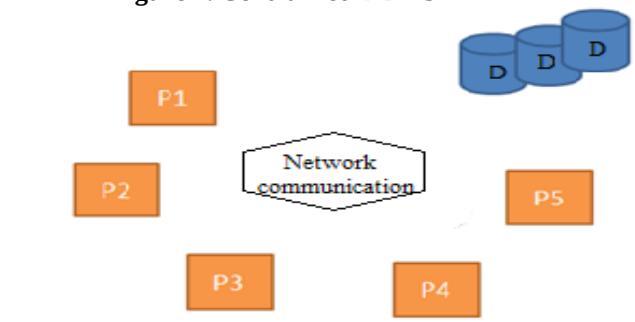
Embedded mobile databases are designed to execute query fragments which are transmitted to them. Afterwards, they combine the individual results in accordance with a schedule that is generated by

directory nodes. These directory nodes received the query before returning the result. Various elements of architecture (applicant nodes (AN), directory nodes (DN) and database nodes (DBN)) swap messages between them in order to execute the submitted queries in the context of distributed database and produce results. (Frank, 2010)

Distributed databases - are a collection of interconnected logical and physical data distibuited in a network. The usual database management system is a distributed database management software that allows the existence of distributed databases and makes data distribution to be transparent to users.

The benefits of distributed database management systems reflect the organizational structure of data, local autonomy, improved security, distributed transaction performance, transparent management of distributed, fragmented and replicated data. Transparency requires the separation of the upper level from the lower level of system implementation errors. The biggest challenge in a distributed system is to provide independent data in a distributed environment: transparent network (distribution), replication transparency, fragmentation transparency - horizontal (selection), vertical (design). (Pathan, 2011)

Figure 1. Centralized DBMS

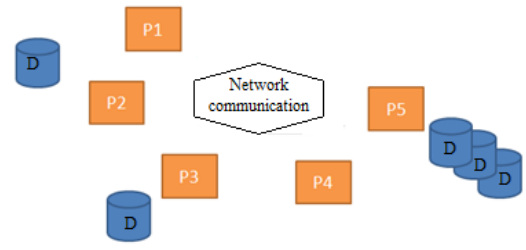


Centralized database management system, in a network

Source (Author)

A comparison between distributed systems and centralized system highlights a number of advantages of the first ones over the latter. Among these advantages there are: the possibility of continous communication between users, share of resources that are available anywhere and which are safe in operations. All these are possible due to the existence of multiple processors. They allow continuous operation of the entire process, even when one of the processors fall. All its tasks will be took over by the other nodes. Other advantages include the extensibility which is a characteristic of distributed systems and also the easy manner in which system updates are done, due to distributed system modularity.

Figure 2. Distributed DBMS



Distributed environment for database management system

Source (Author)

Nowadays different attempts are conducted in order to bridge the gap between two aspects: real-time and distributed computing, where processing, storage, and networking need to be combined and delivered with guaranteed levels of services. The most challenging aspect is the integration of real-time attributes into all levels of distributedsystems. Major focus is given to the need for methodologies, tools, and architectures for complex distributed systems that address the practical issues of performance guarantees, timed execution, real-time management of resources, synchronized communication under various load conditions. (Kyriazis, 2011)

Mobile databases are portable databases that are physically separated from the central database server, but they are able to communicate with the remote server, allowing data sharing. Functionalities required by the mobile databases management system include: communication with the centralized database server, through the nodes, wireless networks and internet access, data replication on the centralized database server and on the mobile device, data synchronization from the central server with the mobile device, data capturing from different sources using the internet, analyzing and manipulating data from the portable device.

The main question regarding mobile databases in distributed environment is how to store data in a consistent way across different platforms like computer centers, web servers, and/or mobile computers. This may be important in system integration with more than one legacy/ERP system, e-governance, logistic management, knowledge management, integration of OLPT and OLAP systems etc. The ACID (Atomicity, Consistency, Isolation and Durability) properties of a database are properties delivered by a database management system to make database recovery easier and make it possible in a multi user environment to give concurrent transactions a consistent view of the data in the database. (Frank, 2010)

Mobile applications as structures in distributed environment require the client applications to run in the context of client-server applications.

Some mobile applications require no permanent connection to a network. There are different ways in which a mobile application can connect to a server. An mobile application can run in the absence of a server, without a network connection, but there are cases where the network connection is indispensable, the fully functioning of the application not being affected. The mobile applications that require permanent connection to the network have a server whose presence is critical.

The architecture of mobile applications is on three levels:

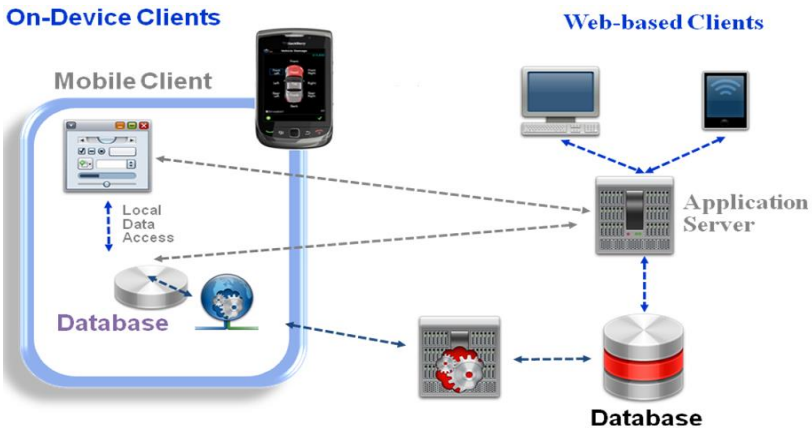
1. The first tier contains the client application running on a mobile device with graphical interface.
2. The second tier consists of applications running on different servers – including servers for mobile application.
3. The third tier is represented by the mobile databases and the storaged data.

3. The benefits of the proposed solution for mobile applications optimization

The proposal of optimizatization of the solutions offered by mobile applications is found in the context of distributed environment. On this purpose the below components of a mobile data management architecture are presented.

- A secure distributed environment and the applications
- Web services-based application server
- The application and data management platform
- Security, monitoring and auditing infrastructure

Figure 3 Ahitecture of mobile applications



Source (Author)

The proposed solution is followed by the software inclusion. Software components are displaced on operating system platforms such as: Linux, Unix and Windows. The software components of the

proposed solution need to undergo a minimum of customization. All applications are web-based, on several levels, with a high degree of availability, scalability and security mechanism.

The secure distributed environment and the application have more users operating in different roles. All users operate independently with different client devices (workstations, laptops ...). All are connected through the communication network.

In a distributed environment of the application there is a distributor of the environment with the following components:

- a client device from the distributed environment, which includes workstations, portables
- local application running on the environment's distributor
- local databases for data storage

The application must be able to work both in online and offline mode. In the offline mode, the application will store the transactions in the local database and it will also access the central server through the web server.

Because the context of intermittent connection might occur, the application must be able to operate without connections. The application must connect automatically, through the availability of local or central connection.

Since multiple users operate independently, the application must synchronize changes made by a user on other users.,

Building a web application that works both in online and in offline mode implies to have both web technologies for the client and browsers: Java applets, Flash ... and also web server application installed in the distributor's location.

Options in choosing browsers support are highlighted in the table below:

Tabel 1. Mobile Browser Client

Device	**Version**
•iPhone / iPad	•All
•RIM BlackBerry	•All
•Google Android	•All
•Microsoft Windows Mobile	•v5, v6 and above
•Symbian	•v9 and above
•Java ME	•All
•Opera Mobile	•All

Advantages of mobile browsers are: eliminating the possibility of hard work to implement infrastructure for any mobile device and there is a greater possibility of application components re-use.

Browser-based mobile BI applications offer unique advantages: portability, data consistency and security. The application and data are on a server, which suppose that nothing is downloaded to the device. This allows that applications and data not to get out of sync, eliminating the propagation of analytic silos. Additionally, it is easier to enforce security since all data, code and passwords are managed centrally on the server.

In terms of database management platform for mobile systems, active market platform is formed by the offerings of several competing companies, such as Sybase, SAP, IBM, Antenna Software, Pyxis Mobile, Apple iPhone, Android. An analysis of the weaknesses and strenghts of these platforms show that there are some challenges in their usage for mobile aplications.

There are limitations in terms of common platform for mobile applications. Efforts to develop platforms are consistent. Also, it is difficult to always have right tools and development environments that correspond to various mobile devices and channels. A challenge is represented even by the programming languages that led to the development paradigm: Java (ME or SE), C, C + + and by the struggles to rebuild services for applications, in terms of logic, visualization and technology used.

The database should have the following characteristics: to support ODBC, JDBC, ADO.NET and to ensure maximum flexibility in choosing application platform.

4. Conclusions

Technologies and mobile devices have evolved in human benefit and also added value to businesses. Nowadays, organizations are not considering anymore the investment in business intelligence elements on mobile systems, just for executives usage, due to the fact an impressive variety of employees need to have data access when they are not at their jobs. Mobile applications have become effective, because they are easy to use on smartphones, tablets and fast networks. The benefits of building up mobile application architectures and successfully managing mobile systems are in terms of possesing the latest information to make strategic decisions, monitoring operational processes and coordinating activities in real time, as waiting is not anymore an option.

The future of mobile business intelligence rely on distributed computing that leverages processing power on the mobile device and remote servers. For the effective competition in the marketplace, mobile business intelligence providers are embracing hybrid architectures to deliver the best of localized and centralized computing with few disadvantages.

5. References:

- *Al-Sakib Khan Pathan; Mukaddim Pathan; Hae Young Lee (2011),* "Advancements in Distributed Computing and Internet Technologies"
- *Bara, A; Botha, I.; Diaconita, V.; Lungu, I; Velicanu, A; Velicanu, M, (2009),* "A model for Business Intelligence Systems' Development*", Informatica Economica Journal, ISSN 1453-1305, Volume 13, Issue 4, pp. 99 - 108*
- *Dimosthenis Kyriazis; Theodora Varvarigou; Kleopatra Konstanteli (July 31, 2011),* "Achieving Real-Time in Distributed Computing for Asymmetric Communication environments," *Proc. of the ACM SIGMOD Conference, California*
- *Gartner (2011) "*Analysis on Sales of Mobile Devices", Grew, *Egham, UK, November 15, 2011— Free Press*
- *Imielinski, T and B.R. Badrinath (2003), Department of Computer Science, Rutgers University, New Brunswick,* "Querying in highly mobile distributed environments"
- Jayaraman, P.P., Zaslavsky, A., Delsing, J, (2010), *"Intelligent Mobile Data Mules for Cost-Efficient Sensor Data Collection"*. In International Journal of Next-Generation Computing, Volume 1, Issue 1, pag.73-90.
- Lars Frank (February 22, 2010), "*Design of Distributed Integrated Heterogeneous or Mobile Databases: System Integration by using relaxed ACID properties"*, LAP Lambert Academic Publishing
- Manole VELICANU; Daniela LITAN; Aura-Mihaela MOCANU (VIRGOLICI), (2010), *"Some Considerations about Modern Database Machines",* Informatica Economica Journal, ISSN 1453-1305, Volume 14, Issue 2, pp. 37 - 44
- Sanjay Madria, Bharat Bhargava, Mukesh Mohania, Sourav Bhowmick Department of Computer Science, University of Missouri-Rolla, USA; Department of Computer Sciences, Purdue University, USA; Department of Computer Science, Western Michigan University, USA; School of Computer Engineering, Nanyang Technological University, Singapore,(2007), *"DATA AND TRANSACTION MANAGEMENT IN A MOBILE ENVIRONMENT",* Chapter 8, pag. 168-184
- *Senthil Kumar, A.V.; Hakikur Rahman (2012)* - "Mobile Computing Techniques in Emerging Markets"
- Xu Yang, Athman Bouguettaya (2010), *Access to Mobile Services (Advances in Database Systems),* Springer Science and Business Media LLC, pag 3-5
- The International Telecommunication Union (2011), http://www.itu.int/ITU-D/ict/facts/2011/material/ICTFactsFigures2011.pdf

CLOUD COMPUTING SECURITY

RADUT Carmen
Associate professor / Ph.D., Faculty of Management Marketing in Economic Affairs , "Constantin Brancoveanu" University, Ramnicu Valcea, Romania, c_radut@yahoo.com

POPA Ionela
Assistant professor / Ph.D., Faculty of Management Marketing in Economic Affairs , "Constantin Brancoveanu" University, Ramnicu Valcea, Romania, popaionela80@yahoo.com

CODREANU Diana
Teaching assistan /Ph.D. student, Faculty of Management Marketing in Economic Affairs , "Constantin Brancoveanu" University, Ramnicu Valcea, Romania, codreanudia@yahoo.com

Abstract: *Cloud Computing implies that your customer information is exchanged via the internet to qualify for various web services, and involves a serious danger in terms of security. Data on the Internet are highly susceptible, while safer when stored in Home/Office on storage media. It is possible to corrupt irreparably, sensitive information, caused by the death of servers, or the worst of an entire data center. Failure prevention measures at the level of suppliers of services related to Cloud Computing presume, interconnection, data encryption and network servers servers for periodical backups of the same files on multiple machines.*

Key words: Cloud, Computing, Security

JEL classification: C88, D85, M15

1. Introduction

Cloud computing is a distributed architecture designed to centralize a server's resources on a scalable platform in order to offer "on-demand" computing power. A true cloud-platform will allocate new storage spaces, extra bandwidth and extra processing power, when needed.

The purpose of cloud computing systems is to reduce administrative tasks and to offer a dynamic environment for the traditional servers.

The use of "cloud" is often extended to refer to any client-server application such as an e-mail client or the Apple iTunes. Nevertheless, a design calculation cloud is based on a scalable architecture which can be used by system engineers.

Cloud Computing is a general term used for anything that involves the supply of internet hosted services. These are generally divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

The "Calculation Cloud" name was inspired by the cloud logo which is often used in order to be represented on the internet.

The purpose of Cloud Computing is to offer the users the freedom of receiving and delivering resources on demand, to configure them as they wish and to pay for the service according to the level of use.

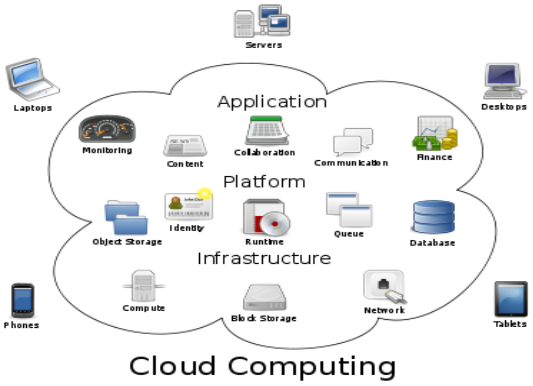
From the service supplier's perspective, both internal and external, cloud computing enables the administration of resources and applications in a multiple-tenancy environment, providing at the same time the personalization and pay per-use. A cloud type of service can represent both infrastructure for application or database hosting and a platform or even an application which can be used on demand.

The word "cloud" as in Cloud Computing means that the architecture takes the shape of a cloud, being easily accessible by users anywhere in the world, on demand. Cloud computing has important advantages in management and it also helps in cutting down the costs. This infrastructure is created by data centers which are being monitored and maintained by the content suppliers. Cloud computing is an extension of this idea in which the business applications are offered as sophisticated services which can be accessed through a network. The service suppliers are stimulated by the profits as a result of attracting clients to use and access the services.

Users such as companies are attracted by the opportunity to cut down or even eliminate the costs associated to "in-house" delivery of these services. This clearly demonstrates the utility of cloud-

computing in terms of business. This obvious identification of demand and offer lead to the creation of cloud computing as a unique discipline in informatics, communication and technology domain (Figure 1).

Figure 1: Main subjects in Cloud Computing

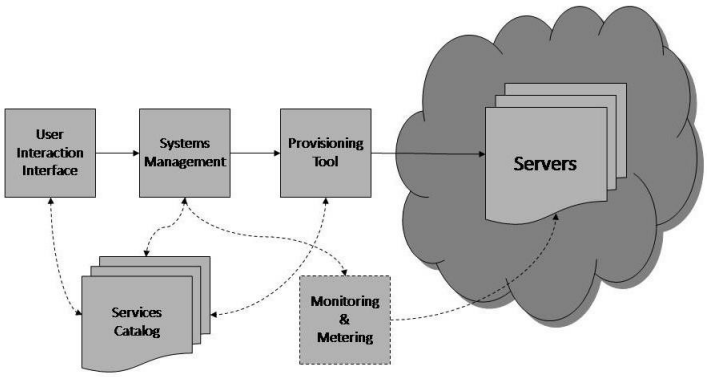


Source: www.fr.wikipedia.org

The success of cloud computing is largely based upon the effective application of its architecture. The architecture of cloud computing is not entirely based on the way it will work with the users.

Cloud computing needs a complex interaction with the hardware which is essential to be able to offer up-time (figure 2).

Figure 2: Cloud Computing architecture



Source: www.newhost.ro

When we talk about a cloud computing system, it is useful to share it in two sections: Front part and Rear part. These two are connected through a network, usually via Internet. The front part is the part which "sees" the computer user. The rear part is the actual cloud of the system.

The front part includes the computer user (or computer network) and the application needed in order to access the cloud computing system. Not all cloud computing systems have the same interface. The services such as web-based e-mail are leverage-programs for the already existent Web browsers such as Internet Explorer or Mozilla Firefox. Other systems have unique applications designed to offer access to the network for the users.

The rear part of the system contains computers, servers and data-storage systems which create the cloud calculation cloud. In theory, a cloud computing system can contain, practically, any computer program which can vary from data processing to video games. Usually, every application will have its own dedicated space on the server.

A central server administers the system, monitors the traffic demands and client in order to make sure that everything is running smoothly. This leads to creating a set of rules called protocols which make use of a special kind of software, called middleware. Middleware allows computers connected to a network to communicate with each other. Most times, servers do not run at maximum capacity. This means that there is unused processing power. It is possible to "trick" a server into thinking that there are

multiple servers, each one running on its own operating system. This technique is also known as "server virtualization".

If a cloud computing company has many users there is the possibility that more storage space is demanded. Some companies need hundreds of digital storage devices. Cloud computing systems need at least to double the hardware usually needed because from time to time a pause is needed. A cloud computing system needs to copy all the information provided by the user and store these copies onto different storage devices. These copies allow the central server to access back-up resources in order to find data which otherwise would be unreachable. Creating copies of the data stored, in order to create a back-up, is called dismissal.

2. Cloud Computing Security

Cloud computing is above all a paradigm change. It is not a new concept, because even in IT, the history is cyclical. The current paradigm is completely different from the mainframes created in the 70's or 80's and even though it is cyclical, history is evolutionary. Cloud computing is not a hype anymore, not even a buzz-word. It makes sense economically speaking and this will soon lead to a large scale application of this paradigm. It is just a matter of time.

The thing that changed in the last decade is the fact that today's companies, smaller or bigger, build its business processes based on informational infrastructures. Simplifying, any business process, ultimately means information. In the current context, this information presents itself in an electronic shape. Cloud computing means moving this information from a predictable and easy-to-control space (local network), in an agile, mobile, less palpable space.

Based on these premises this leads to the main obstacle in the way of large scale adoption of cloud systems: loss of palpable control over the company's essential information.

The main problem is one of trust. Once again, the company's information is critical – it can generate a competitive advantage, development, safety, etc. The things that a client looks for at a cloud provider is being comfortable with delegating control, stability and credibility. Stability is boasted by the prestige, presence and market share of the supplier. Credibility is boasted by its activity history, prestige, professionalism, and the last but not least the accreditations and certifications. Trust is, in the end, a transitive concept.

A second interesting problem is electronic identity. Identity offers the client two basic elements in information manipulation. Authenticating and granting access. In the classical paradigm, identity is offered by the company itself. It is the one which provides its members with the needed credentials. In cloud computing context, the cloud computing supplier is also an identity provider. The control for identity management is delegated.

In this context, important is the actual owner of the identity. The cloud supplier must grant absolute control to the company over its identity elements.

The third fundamental problem is the one of actual control over the company's information. How can the owner control who, what, how and where is the information accessed since the information is not stored locally but on a remote location. Traditional information security inside a company relies on an obvious constant: the size of the company's perimeter. In the cloud computing context, this perimeter is not well defined anymore. At this point, the cloud supplier must offer enough security control in order to compensate.

Now, the suppliers of security software noticed the evolution of cloud computing. Cloud is a reality and the transition already started. Terms such as "private cloud" or "hybrid cloud" are only intermediate stages to a global cloud.

Obviously, the threats, risks and vulnerabilities are also evolving. The impact which a cyber-attack can have at this moment is growing logarithmically. A completely informational and interconnected and globalized world is nothing but a bigger cyber-attack surface from a hacker's point of view. In this context traditional security checks adapt and evolve also. Virtualized Firewall/VPN, Virtualized IDS/IPS, Hybrid Content Filtering, Hybrid Data Loss Prevention are actually traditional security measures adapted to the new model of information representation and they are merely intermediate stages to a "anything-as-a-service" concept which itself is nothing new.

Other problems related to cloud computing service are regarding legal aspects. Besides the security issues mentioned above, the cloud computing service suppliers and the clients will have to negotiate the terms around the level of responsibility (regarding data loss and the way these issues will be addressed), intellectual property and end-of service (when the information stored will return to the client).

Legal issues may include, also, record keeping of the public sector demands, if more agencies are obligated by law to keep and offer electronic recordings in a specific way. This aspect can be determined by legislation, or the law can make the agencies follow the rules and practices established by a recording-keeping company. Public agencies which use the cloud computing service for data storage, must take into consideration these issues.

3. Conclusions

Cloud Computing has a strong impact on the IT industry, thanks to technological innovations to support the new trends in the world of software. Analysts have found that organizations are in a period of transition from hardware to software owned by the company from a model which involves their use as services of payment depending on use. Cloud computing allows to increase efficiency without affecting the IT market.

4. References

- Christopher M. Moyer, Building applications in the cloud : concepts, patterns, and objects, Addison-Wesley Professional, 2011
- John W. Rittinghouse, James F. Ransome, Cloud computing : implementation, management, and security, CRC PRESS, 2010
- George Reese, Cloud application architectures : [building applications and infrastructure in the cloud], Publisher: O'Reilly Media, 2009
- http://en.wikipedia.org/wiki/Cloud_computing
- www.newhost.ro
- www.fr.wikipedia.org

SECURITY SOLUTIONS FOR DATA AT REST

ŞERBAN Mariuţa
Ph. D. Student, Department of Computer Science, University of Piteşti, Piteşti, t_mariuta@yahoo.com

ŞTEFAN Raluca-Mariana
Ph. D. Student, Department Economic Informatics, Academy of Economic Studies, Bucharest, stefan2012@yahoo.com

Abstract: *Due to the rapid development of information technology and of the Internet, most companies keep their data stored in a digital format, in databases. Database systems have become an essential component of everyday life in the modern society, as most people perform activities that assume interaction with a database, such as: banking, reservations of any kind, searching for a book in a library or in a bookstore.This paper describes a new approach regarding the encryption for protecting the data at rest from databases*.

Key words: *database, encryption, cryptography, encryption algorithms*

JEL classification: C 61, C 63, C 88, P 41

1. Introduction

Cryptography can be defined as the science of secret writing and it uses mathematical methods for data transformation in order to hide their content or to protect them against changes.
Encryption, the process of disguising data so that its substance is hidden, is a very effective way to achieve security for data. Database encryption strategy raises several important factors that must be taken into consideration, for example: when the encryption is performed inside the database engine or elsewhere in the application the question is if the database is still safe.

In order to make the information impossible to read by an individual who does not hold the key, cryptography uses an algorithm. The key is used to lock or unlock the data. Encryption is used to protect network data, in systems where simple authorization is not sufficient to protect data or as an extreme measure of data protection.

The concern of cryptography has been dedicated to strengthening ciphers by complicating algorithms through combining substitutions and transpositions of symbols or blocks.

What worth noting is the fact that this science serves two worlds with different interests, the first is the world of communications, such as those between the rightful users of the same database, representing legal operations and the second one is the world of illegal operations in which the unauthorized user tries to intercept or alter messages, sometimes even to destroy them. From this circumstance, cryptography appears as a continuous struggle between the two parties. The success of the unauthorized person to enter the database will imminently lead to strengthening the safety measures undertaken by the authorized users.

This paper presents the functionality of cryptography and discusses how this applies to authentication and encryption, presenting examples of ways to use cryptography in Microsoft SQL Server by using the technologies and tools provided by this server, technologies that are provided by the majority of competitors.

2. Database encryption in Microsoft SQL Server

The tools for data encryption provided by Microsoft SQL Server use a variety of algorithms and encryption keys that can be operated at database and server level, thus providing maximum transparency for applications that connect to them. Key management can be achieved even at user level, so the user has the possibility to keep data secret, data that the administrator cannot have access to.

When data encryption and decryption is carried out at database server level then the data circulating on the network will be clear and will be exposed to attacks. In a SQL Server you can secure the transmission channel by using certificates. There is the possibility to automatically generate a certificate used for channel authentication, if there isn't installed a certificate that works explicitly with a SQL Server.

Encrypting data at rest is done inside of the database. The majority of attacks do not happen on data in motion, but at the end point of them, where data can rest for large periods of time.

If an organization decides to encrypt the data at rest, it must also take a decision on the granular level of encrypted data.

Once the transparent data encryption was introduced, users have the opportunity of choosing between cell level encryption – with SQL Server, database level encryption – using TDE, and file encryption options – offered by Windows.

TDE represents the best choice for encryption due to its maximum compliance with regulations or with company standards of security data.

TDE works at the file level is similar to the two features of Windows: The file encryption (EFS-Encrypting File System) and BitLocker Drive Encryption; however EFS and BitLocker cannot be replaced by TDE.

2.1. Encryption at database level

Encryption at database level generates a single file that represents the entire database. An advantage of this level of encryption is the security of all data. This great advantage turns out to be also a disadvantage because whenever you want to read or write to or from the database it must be entirely decrypted.

Transparent data encryption (Transparent Data Encryption - TDE) is an encryption feature introduced in Microsoft SQL Server 2008. This feature is designed to provide protection to the entire database at rest, without affecting existing applications.

The implementation of encryption in a database, traditionally involves complicated changes such as changing the basic schemes, removing functionalities and significant decrease in performance. For example, when using encryption in Microsoft SQL Server 2005, the type of the column data has to be replaced with the varbinary data type and equal type searches are not allowed, the request referring to stored procedures that are automatically used, having to deal with encryption and decryption and performance, all leading to a slow query.

These aspects are not unique to SQL Server, other database management systems facing similar limitations. Customized systems are often used to solve queries of equality and inequality searches, options that cannot be used for all queries of this type. Even general operations of a database, such as creating an index using foreign keys several times do not work at the cell or column-level of encryption systems because using these features inherently determines information leakage.

Transparent Data Encryption solves these problems by encrypting everything so that all types of data, keys, indexes, and so on can be used to their maximum potential without jeopardizing the security or leakage of information from the storage devices.

Transparent Data Encryption (TDE) is a technology used by Microsoft and Oracle to encrypt the contents of a database. TDE solves the problem of protecting data at rest through the encryption of the database stored on different storage devices. Companies often use TDE to address compliance issues. TDE is provided by Microsoft as part of Microsoft SQL Server.

Microsoft SQL Server provides two levels of encryption: encrypting at database level and encrypting at cell level. Both levels of encryption use the hierarchy of keys management.

2.2. Encryption at row or column level

Encryption at row or column level represents ciphering a row or a column from within a table. You need to carefully choose which column will be encrypted in order not to impede the operations that are performed in the database. It is avoided to encrypt a column that refers to a <WHERE> clause or that represents a primary key column because this requires column decryption prior to any comparisons.

Cell level encryption was presented in Microsoft SQL Server 2005. Cell-level encryption is implemented by using a series of built-ins and the hierarchy of key management. Using this type of encryption represents a process that requires a rearrangement of the call requests for the encryption and decryption functions. In addition, the schemes need to be modified in order to store varbinary data type data and then at the reading these should be returned to the appropriate data type. The traditional encryption limitations are inherent for this method because none of the techniques of automatic query optimization can be used.

2.3. Using cell level encryption together with TDE

Cell level encryption can be used for in depth defense for a database encrypted with TDE technology as well as for constraining access control using passwords.

With all the disadvantages of using cell-level encryption, by using TDE, cell-level encryption can be useful for a subset of extremely sensitive data.

Usually, TDE and cell-level encryption meet two different objectives. If the amount of data to be encrypted is very small or if the request can be custom designed and the performance level is not a concern, cell level encryption is recommended over TDE.

Otherwise, TDE encryption is recommended for encrypting existing high performance applications or for sensitive applications.

3. Security at operating system level

An advantage for users who need password or private keys based protection for applications in a database is provided by operating systems via an API (Application-Programming Interface) data protection system.

One of the functions of this system is DPAPI (Data Protection API). This manner of protecting data is a service that ensures data privacy through encryption. Thus any application can secure its data without needing a cryptographic code, besides calling the necessary functions of DPAPI.

Data protection API provides password based protection and DPAPI uses the Triple-DES algorithm together with strong keys.

3.1. Encrypting File System (EFS)

Encrypting File System – the files encryption system that was introduced in Windows 2000 – is based on Windows API Cryptography (CAPI). Both files and folders can be marked as encrypted, even though encryption takes place only at file level. Each file is encrypted with an individual file encryption key (FEK), more than that each database is encrypted with an individual DEK key in TDE. FEK is protected by a user's certificate, similar to how DEK is protected by a certificate.

The EFS certificate is assigned to a user, while the TDE certificate is conceptually awarded server-wide to an object. To encrypt FEK there can be used more certificates, which allow multiple users to access a file. When using EFS with SQL Server, the duty account of the database server must have access to the file containing the encryption keys, so that through the encryption of any file from the database, that file can be read. This cannot be used as a form of access control – the duty account is used to read the database files, regardless of the login account.

EFS, encryption at operating system level presents some advantages; because TDE is a technique limited to database files, EFS allows non-basic data encryption and even file level encryption, which assumes larger coverage of the encryption area.

Key management is abstracted for the operating system, which allows users to access Windows certificate storage.

Encrypting File System offers a way to recover data when keys are lost, unlike TDE which currently does not have a similar solution.

The disadvantages of EFS over TDE are mainly related to performance and management.

EFS requires file administration privileges at operating system level, which the database administrator (DBA) could not have. Because protection is related to EFS, database deployment, database backup or adding groups of files cannot be protected if these files are stored in other locations unprotected by EFS.

Also, EFS implementation can be different. Normally, this is not a problem because EFS is used mainly on a single computer for only one user, but it should still be considered.

EFS is best to be used on laptops, computers, workstations where the database is accessed by a small number of users.

Because of performance concerns, it is generally not recommended to use EFS together with TDE, though nothing prohibits this. Using EFS with TDE is a viable option when performance is not a priority and when an in depth defense is needed. EFS can be used instead of TDE for granularity at file group level. Also, because EFS protects data at file level, it can be used to protect partitions where data is temporarily written to disk. In environments where performance is a major concern, EFS is not recommended to be used together with a SQL Server.

3.2. BitLocker

BitLocker is a technique of volume encryption that is included in various operating systems like Windows Vista Enterprise Edition.

Both BitLocker and TDE protect against offline attacks. BitLocker protects at volume level so that when the server is online, the volume is unlocked, but not decrypted. BitLocker, like EFS, has a mechanism of data recovery that TDE does have yet.

The advantages of BitLocker are ease of administration, abstraction management and transparency. The disadvantage is that the protection extends only to the volume. The detaching or backup copy of the database to another volume, which is not protected by EFS or BitLocker can lead to a decrease in file protection, leading to the possibility of even losing the file. Another disadvantage of BitLocker is the broad area of protection. Because the entire volume is unlocked, any user with access to the computer can access the files on the disk and can also access data from the original text.

Similar to the TDE, BitLocker is based on other access control mechanisms such as permissions to use the database in TDE and permissions for BitLocker to use Windows files. As with EFS, the database administrator may not have the required privileges in order to manage BitLocker.

BitLocker does not have associated the same performance concerns as EFS, which is why it is recommended to use BitLocker together with TDE for increased security.

BitLocker and TDE encryption technique provides minimum performance data encryption and requires additional management hours. Any user who has the right to view tables can see the data unencrypted plain text.

Cell level encryption remaining encrypted in the memory, without having access to the key and to the explicit description of the data, users that have access to table reading will be able to see the encrypted information. Unlike TDE, this function is available in SQL Server 2005 in all its editions.

Thus, for performance reasons, a pair of public/private keys (asymmetric or certified keys) designed for data encryption and decryption should not be customary used. Performance is very poor when using an asymmetric key encryption algorithm. The real advantage of using pairs of keys is that any user can encrypt, but only the private key owner can decrypt. This means that the asymmetric and certified keys can solve the problem of communication between the sender and the receiver of the information - problem that does not always apply to the database engine. For a database system such as SQL Server, encryption and decryption are performed by the same agent (engine or service agent, thus becoming meaningless to use the asymmetric keys encryption algorithm and the result of using it leading to a weak and unnecessary performance).

Certificates usage is preferred in cases where encryption or decryption is not administered locally or it has to be managed by different users. The best security solution involves encrypting data with a random symmetric key and the subsequent use of the certificate to encrypt this encryption key. The classes of .NET symmetric encryption keys have the ability to generate random symmetric keys.

SQL Server encrypted data must be stored in a varbinary column and the value encrypted in this matters cannot exceed 8000 bytes. Because the encrypted values tend to be higher than the original text, the maximum size of data encryption from the considered text should be less than 8000 bytes. The impact of encryption on the storage capacity varies depending on the used algorithm.

It is possible to use the size of encrypted data to infer information about the data from the original text. Even if the encryption algorithms have fixed category sizes it is not possible to assess the size of the encrypted values in order to get clues about the size value of the original text.

4. Quantitative testing for two methods of encryption with SQL Server

A comparison is made for two applied methods of encryption namely encryption using TDE and column level encryption.

Enabling TDE database will be marked as encrypted, and the state of DEK (Database Encryption Key - DEK is a symmetric key secured by a certificate that is stored in the master database or asymmetric key protected by the EKM module - Encryption Key Management-how the symmetric and asymmetric keys stored outside SQL Server) is set as "Encryption in Progress". This server starts a thread (called encryption scan) in the background that explores all the database files by encrypting them (or decrypting them if TDE is disabled). This line creates a shared database lock. Even when the DDL executions take place, an update lock shared database is generated asynchronously running thread that performs encryption using shared lock.

The only operations that are not allowed when encryption is the process are those that change the file structure and those that put the database offline by posting them. Exploration is running over the virtual file that stores transaction log so that future writings in the file are also encrypted.

When data base encryption is over, state of DEK is encrypted. Encryption algorithms bear are AES keys up to 256 bit or triple DES. Database file size is not altered by encryption as opposed to save your file size transactions, size becoming larger.

Enabling TDE are encrypted all data found in the database file, files which is stored in log files transactions, even back-up data from disk. Because TDE does not protect data in transit, then the data in use are not encrypted.

When activating TDE for a database is automatically activated also temporary database encryption (tempdb).

The following stages are necessary to implement TDE:

1. It creates a master key as follows:

```
use Cards
create master key encryption by password='rate'
use Cards
alter master key
add encryption by service master key
```

2. It creates a protected certificate by master key created previously

```
use Cards
create certificate TDECert
with subject ='TDE Certificate'
```

3. It creates an database encryption key (DEK) protected by the above mentioned certificate:

```
USE Cards
CREATE DATABASE ENCRYPTION KEY
WITH ALGORITHM = AES_256
```

4. Enable TDE:

```
ALTER DATABASE Cards
SET ENCRYPTION ON
```

For data storage has implemented a database called test, with 87500 records with which to test table Values encryption using TDE and comparison execution time of the instructions SELECT ALL, SELECT RANGE, SELECT MATCH.

For quantitative testing of the column encryption level is implemented another database called test2 with the 87500 records. In the table Values is created a new column called Encrypted_Sum which will be stored encrypted values from column Sum, aiming the execution time for same instruction: SELECT ALL, SELECT RANGE, SELECT MATCH.

Quantitative comparison of time required for two encryption methods tested above and the time required for different types of searches is illustrated in the following table:

Table 1: Time of during different types of searches

No. records	Encryption method	Type of search	Time/seconds
87500	TDE	SELECT ALL	4
		SELECT RANGE	1
		SELECT MATCH	0
87500	Column Encryption	SELECT ALL	5
		SELECT RANGE	0
		SELECT MATCH	0

5. Conclusions

The encryption solutions desired by database users present more options in SQL Server and Windows. These options are not mutually exclusive. The different levels of encryption available in SQL Server and Windows can be used to trigger the defense system in depth and to globally secure the database.

EFS and BitLocker are valid solutions, either together with TDE or as standalone encryption systems, either on their own; TDE is not designed to replace these solutions.

BitLocker, EFS and, in a much more limited extent, cell level encryption can be used to protect database systems (master, model, of resources and msdb) which currently cannot be encrypted by TDE.

The most common problem related to the functionality of cryptography is that it is used to address and not to solve problems. Situations that are frequently encountered can be grouped into two categories of examples:

- using asymmetric key pairs in situations when communicating the key is not an issue, such as data stored in a SQL Server database;

- using encryption for data access control. In SQL Server permissions are robust and designed for this. The only exception occurs when access to the mass-media storage may be compromised, encryption adding a significant level of security.

Although SQL Server provides powerful tools for encryption and verification, most often these are not understood well enough, which can lead to a poor or incomplete implementation.

6. References

- Barnes, R. (2011) *Database Security and Auditing: Leading Practices*, Enterprise Auditing Solutions Applications Security.
- Băjenescu, T.I. (2003) Progresele informaticii, criptografiei şi telecomunicaţiilor în secolul 20, Editura Matrix Rom, Bucureşti.
- Cocianu, C.L.; State, L.; Uscatu C.R.; Ştefan R.M. (2011) Learning from Data – A SVM Based Approach, Economy Informatics 11 (1), p. 5-14
- Crampton, J. (2009) Cryptographically-enforced hierarchical access control with multiple keys, The Journal of Logic and Algebraic Programming 78, p. 690-700.
- Fusaru, D. (2002) Arhitectura bazelor de date. Mediul SQL, Editura Fundaţiei România de Mâine, Bucureşti.
- Fusaru, D.; Şerban, M. (2010), Algoritmi de criptare a bazelor de date, Annals of Spiru Hart University, Economic Series, 1(10), p. 207-215.
- Hicks, J. (2008) Criptography in SQL Server.
- http://msdn.microsoft.com/en-us/library/cc837966%28v=sql.100%29.aspx
- Hsueh, S. (2008), Database encryption in SQL Server 2008, Enterprise Edition, February.
- http://msdn.microsoft.com/en-us/library/cc278098(v=sql.100).aspx
- Lesov, P. (2008), Database Security: A Historical Perspective, University of Minnesota, CS 8701.
- Patriciu, V.V.; Ene-Pietroşanu, M.; Bica, I.; Văduva, C.; Voicu, N. (2001), Securitatea comerţului electronic, Editura All, Bucureşti.
- Srikanth, Radhakrishna (2011) Database security best practices, www.helium.com.
- http://en.wikipedia.org/wiki/Database_security.

EMERGING IT TECHNOLOGIES – ADVANTAGES AND RISKS

STANCIU Victoria
Professor, Ph.D, Faculty of Accounting and Management Information Systems, Academy of Economic Studies, Bucharest, Romania, victoria.stanciu@cig.ase.ro

BRAN Florin
Economist Ph. D, BCR Erste, Bucharest, Romania

Abstract: *Companies are looking for new business models ensuring quality, effectiveness, efficiency and competitive advantages. IT-based business models implementing the most recent IT developments are considered the new vector in business innovation and key of success. Companies are aware that the IT specific risks can induce failure in achieving strategic objectives. Managing risks, of all kind, is the key of success. The risks must be monitored and the new threats arising from the new IT solutions must find their correspondent in security policies and solutions.The present paper aims at emphasizing the importance of IT in the business development and management signaling the new emerging risks companies must face when shifting their business on new IT environments and models. The empirical research has focused on the IT security policies' adjustment to the new IT developments imbedded in the businesses' processes.*

Key words: IT risks, cyber-attacks, risk management, IT strategy.

JEL classification: M15, G30, D83.

1. Introduction

As a result of the innovations in business models and the permanent updates in IT infrastructure aiming at developing and strengthen the companies' activities new emerging risk are registered. The companies are more and more aware that information technology must be treated in its dual approach: the competitive advantage brought in business development and market recognition and emerging risks determined by the technologic development. Managing information and IT risks is imperious necessary for the business stability, growth and company's reputation. Security gaps can expose the company to unforeseeable legal, financial and reputation exposure. Security management focus on compliance requirements is not longer responding to the risks' landscape and business's needs. That is why companies are concern in implementing a solid IT risk management process addressing the large scale and continuous increasing landscape of risks. An effective risk management process is based on IT assets' assessment indicating the companies' IT risk exposure. This evaluation provides the information that allows risk management decision such as risk acceptance, risk mitigation or transfer conferring stability and protection to the business and facilitating the objectives' achievement.

The companies are more and more aware of the importance of strong IT governance imbedded in the company's global governance process. Based on the governance guiding principles and lines are issued strategies that are put in place by effective policies. Business IT-based implies compliance with security standards and good practice guidelines. In this respect, the companies' security policy has to be aligned to the security standards and/or regulations specific for the existing business IT solutions. A good and sustainable security policy has to be flexible, build on security components that are adjustable and easy to assimilate emergent technologies in cost-effective solutions. The companies are more and more aware that maintaining competitive advantage implies a profound understanding of the new IT implemented solutions and their specific risks. The companies must be prepared to respond to the unexpected risks, the black swans too, and this requires a dynamic approach in risk management.

The diversity, inventive nature of cyber-crime attacks and their significant impact determine a solid awareness on IT risks not only in the information security professionals' world but also at top management executive levels understanding the potential affectation of the business they are running. The executives have become aware of how strongly the enterprise relies on IT, the importance of IT dimension in the business strategy and their responsibility on IT risk management. In this respect, security has started to be integrated into decision making process.

New concepts related to IT security have been created to reflect new emerging IT technologies, IT-based business models and cyber attacks. Social networking, consumetisation, security patches, cloud

computing, botnets, MITB ("man-in-the-middle") are just some of the concepts we are dealing with (as business managers or security specialists) trying to understand how the new information technologies can be used and provide significant business increase and trying at ensuring also protection for the IT assets and business itself from the specific risks.

Our research aimed at analyzing the IT risk exposure determined by the new information technologies' assimilation in the businesses' processes and the changes that should be registered in the global governance and management processes with focus on risk management, decision making process, innovation and strategic development.

In their work, the authors opt for a mixed research methodology, including deductive positive research and inductive critical interpretative research. These two approaches ensured a strong interconnection of theory and research work, merging quantitative and qualitative research methods. The methodological approach, followed by the authors, was developed on the above mentioned steps:

- Delimitation of the research area;
- Research area literature review: the main source of information was represented by IT security surveys, new technologies and their business implementation solutions, best practice guides issued by IT professional organizations;
- Delimitation of the conceptual framework boundaries which included the specialists' points of view regarding fundamental concepts and the relationships arising among these concepts. At this level, descriptive, exploratory and evaluative research techniques were used;
- Decisions related to the key areas to be investigated: new IT emerging technologies and the risk raised as a result of their implementation in the companies' information systems;
- Carrying out interviews with IT managers and specialists, documentation on risk management processes implemented in Romanian organizations and conducting analysis on IT risk management solutions.
- Final conclusions formulation.

The conclusions retained based on our research will be discussed in the following chapters aiming at providing insights on IT risks raised from new IT developments, risk management process, IT security policies and strategies.

2. Business and new IT technologies and services

In a competitive world, succeeds the most innovative, flexible and efficient competitor. Companies are looking to find the optimum equilibrium between costs and benefits. Running business today is almost homonym with using emerging IT technologies. IT investments create value by ensuring business increase and diversification, customized offer, improved business processes and risk management. In this respect, companies are changing their business models being focused on IT-based solutions. The implementation of emerging IT solutions determines significant impact on the business itself, on its management and employees who are asked to adapt their work to the new environment ensuring higher productivity and effectiveness. IT is today's business success vector bringing opportunities but facing companies with a new challenge: new emerging risks. Facing IT risks and especially risk related to e-crime imposes good knowledge related to the emerging information technologies, the today' threat landscape and new IT-based business models approached from e-crime risk perspective. In this respect, the present paper focus on the risks emerged from the new IT-based business solutions and the ways to mitigate those risks.

Social networks - fertile space for gathering information

Today's appreciated social networking services seems to be not only a virtual environment designated to build contacts and "meet" people sharing same activities, ideas and interests but also a fertile field used by hackers for gathering precious information related to people. More than that, social networking become the targeted space where some companies are collecting personal data without the user's knowledge or consent and then either sharing it with third parties or simply keeping it for future use. On the social networks people are defining their profile specifying their present and previous professional activity, share information about their hobbies, professional and personal achievements, family events and issues. As the social networking is developing and gathers more and more people, increases the e-crime attackers' chase for "precious" people's information. Being so dependent of our work and wishing to develop our professional horizon and contacts, we are creating professional networks that link colleagues, clients, and specialists in related fields debating projects, professional

issues and critical problems looking for answers and solutions. In this context spear-phishing finds a fertile area to "pick-up" information. Making public information related to them, people become vulnerable.

In business area we face two different opinions. For example, senior executives in marketing area see the tremendous benefits of engaging in the ways that customers prefer. But numerous voices are reluctant to the social media usage inside the company. This second trend emphasizes the risks induced by these IT developments. As a result, global studies states that most organizations recognize that use of social media by employees introduces significant risks and have banned these service (Kaspersky, 2011).

To reconcile the two parts (supporters and opponents of social networking inside the company) it should be carefully analyzed the potential benefits for the business of such services' usage. If there is a solid justification for it, there must be put in place clear and well disseminated policies specifying the role and usage of those services for the business purposes only, delimitating the information that can be disseminated in this space and ensuring the needed protection for the end users' communication.

To BYOD or not BYOD

Companies have understood that global economy, high competition and the changing economical environment impose new business models: innovative, flexible and IT-based business models. In this respect, companies are opened to new solutions ensuring business increase, productivity and operational output. Over the past few years we were witnesses of consumerization phenomenon that significantly impacted the IT culture: users are in the front line of most recent technologies' acquisition. The individuals' satisfaction for their new portable devices determined them to use their own portable devices for business purposes agreeing with the employer to support also part of services' costs just to benefit from the facilities offered by those top devices (Good Technology, 2011). Good Technology State of BYOD Report (Good Technology, 2011) emphasizes that "half of the respondents with BYOD ("bring your own device") devices in place said that employees cover all costs associated with their devices including device and data plans; 25% of respondents use a stipend to encourage participation and help cover costs". As a result BYOD bring important cost reductions for the companies but also generates some advantages over competitors (organizations using BYOD programs benefit of the latest features and capabilities) and the applications that mobilized units and people have registered significant success determining their classification as "critical applications". BYOD programs are also raising significant security problems: sensitive information is stored on employees' personal devices, personal and corporate applications are running on the same personal devices and sensitive information flows in wireless environments being an attractive target for cyber attacks. Embracing BYOD, companies are losing the control over the IT hardware and even on its data. In this respect, clear defined security policies are needed to ensure companies' data and systems security. Such policies should state minimum security requirements for allowing personal devices to connect to the companies' network and databases. The policies have to cover multiple aspects of BYOD as for example data ownership, compliance with specific regulation related to data privacy, retrieval of company's data on the personal portable devices and data destroy when the employee is leaving the company etc.

All above mentioned aspects emphasize that consumerization requires a strategic approach aiming at mitigate security risks, financial and compliance exposure.

On cloud but still in compliance?

Companies are looking for IT solutions ensuring improved resources' utilization and cost savings. In this respect, virtualization and cloud computing is one of the emerging IT solutions that capture attention. Aiming at implementing virtualization and cloud technologies as quick as possible so that to benefit from all the advantages provided, companies seem to take fewer precautions on the security facet of the problem forgetting that traditional physical server security is not addressing in an adequate manner the specific risks raised by these new IT developments. We have in view, for example, the communication blind spots and inter-VM attacks. The key in understanding the particularities of virtual environment security is that if one element of the environment is compromised any other element may also be compromised if virtualization-aware security is not implemented. As for example if one VM is compromised as a result of a cyber-attack any other VM, placed on the same host, can be also compromised or in case of a hyperjacking the compromise VM may attack the hypervisor. Another common target for the attackers is the hypervisor itself, based on the fact that his role is to enable multiple VMs to run within a single computer (TrendMicro, 2011a).

There are multiple attacks scenarios that must be taken into consideration and the security solution stays on the provider's ability and knowledge. In the company's "traditional" IT environment, based on the security policies, different security levels are provided for sensitive data comparing with less sensitive ones. This desiderate is harder to be achieved in cloud platforms as long as critical data and less sensitive data are stored on different VMs on the same host. The solution can be represented by self-defending VM security being insured protection such as intrusion detection and prevention, a firewall, log inspection, and antivirus capabilities.

Our research emphasized some risk exposures emerged from compliance perspective. The most exposed are large companies running global businesses and organizations acting in most regulated industries as for example banking and financing. The compliance risk is determined by the fact that in a cloud environment data migrates from one location to another for optimization purposes for example. In case of regulated data, as for example banking data, this migration is performed passing through different jurisdictions, data owner having no means to control this migration but remaining responsible for compliance issues.

Another aspect related to the security of highly regulated data in cloud is in regard with encryption and decryption processes being known the different regulations on this topic in different countries.

Particular issues arise in relation with banking most sensitive data, particular regulations requesting storage devices' destroy in case of deleting data so that no data recovery to be possible. This requirement is hard to be put in place in case of cloud computing and more than that hard to audit.

And to add another aspect in the landscape of risks and constraints related to cloud computing it can be emphasized a huge problem: how can the auditors (internal and external) make their jobs? Getting evidence of system integrity and compliance in multiple aspects (some being already underline in our presentation) is impossible as a result of technology itself and the data flows which are different from any other "traditional" environment. To all of these, the refuse of cloud providers to allow auditors in their premises to perform their mission ends the whole discussion.

Virtualization and cloud remain a feasible solution for many businesses, but the decision to migrate in this "space" has to be taken based on multiple aspects' analysis – security and compliance issues have to be included.

3. Risk management

Security specialists recognize the increasing sophistication of threats: threats are more ingenious and more difficult to detect. Another characteristic of the IT threat landscape is the increasing availability of dedicated tools on the internet. As a result, we are dealing with an increase number of potential cyber-crime act initiators; cyber crime acts ended to be the privilege of "specialists" being now performed by individuals having limited computer skills. Even so, the diversity and impact of cyber attacks have registered significant increase. As global IT security surveys reveal "significant number of businesses have already become victims to cyber crime, including targeted attacks, events of corporate espionage and loss of sensitive intellectual property. This in turn leads to the conclusion that cyber threats have become much more important for business, which was confirmed by 46% of the organizations" (Kaspersky, 2011). The companies must be aware that is a matter of time to be subject of an IT attack. The question is when and how the security incident will occur. In this respect, the companies must be prepared to have an adequate response at any kind of security events.

The same survey mentioned above states that "in the last 12 months 91% of companies have experienced at least one IT security event from an external source. The most common threat comes in the form of viruses, spyware and other malicious programs. 31% of malware attacks resulted in some form of data loss, with 10% of companies reporting loss of sensitive business data. The second most frequent accident is network intrusion; 44% of companies surveyed experienced a security issue related to vulnerabilities in existing software. 18% of the organizations also reported intentional leaks or data being shared by staff" (Kaspersky, 2011).

IT risk, in many organizations, is still understood as a technical issue and is not approached linked to other specific risks. That is why IT risks are considered to be just on IT specialists' responsibility and there is no adequate implication of other senior executives. As a response to this approach IT Governance Institute (ITGI) document on IT risk framework states that the "responsibility belongs to those who must ensure that the activities are completed successfully. Accountability applies to

those who own the required resources and have the authority to approve the execution and/or accept the outcome of an activity within specific risk IT processes" (ITGI, 2009).

Statistically the most frequent types of external threats are represented by: viruses/worms/spyware, spam, phishing attacks, network intrusion, DDoS, theft of hardware from premises or off-site, corporate espionage. As a result, the companies' reaction is structured on the main following directions:

- Anti-malware protection
- Firewalls
- Improved access policy by implementing levels of access to different IT systems based on granted privileges;
- Physical security for IT systems;
- Improved physical control access policies;
- Data back-ups and recovery policies and plan.

A special chapter of threat landscape raises many concerns in CISOs' world: the Advanced Persistent Threats (APTs) recognized by their ability to remain undetected while the attack is executed. The attacks are developed so that they cannot be conclusively traced. This is the reason why, APTs are referred as being initiated by state-sponsored organizations. The lexicon of APTs includes components as for example Stuxnet and Operation Aurora that produced significant concerns. Stuxnet is a computer worm that targeted Siemens industrial software and equipment. Even if it is not the first time that hackers have targeted industrial systems it is the first discovered malware that has spied on and subverted industrial systems. Remaining in the same register of threat landscape is hard not to mention viruses/Trojans like SpyEye and ZeuS that developed, not long ago, powerful attacks on banks.

The dynamic and increasing impact of IT threats make traditional controls less effective. Where is the solution? Maybe looking above the existing threats and treat risks in a holistic approach rather than establishing individual-risk solutions.

The IT global security surveys emphasized the priorities from security officers' point of view, based on the internal and external findings (Deloitte, 2010):

- identity and access management;
- data protection;
- security infrastructure improvement;
- regulatory and legislative compliance.

Identity and access management represent a milestone of the information security process. Who must be the system's users? How users are identified by the system? Who must access data? What data can be accessed by each user? How and when data can be accessed by different users? These questions, and more others, seem to be simple at the firs look, but the right answer is hard to be provided. First of all there is no unique answer for above questions. Each company, based on its security policy, will provide different answers. Granting access rights to the users is based on a rigorous relationship established between the users' attributions in the company and the data they have to access in order to achieve their responsibilities. Granting the access (type of access) depends also on the company's business model, the existing IT infrastructure, the company's culture and risk awareness.

Protecting data is a complex process implying: data classification, granting access rights to the users, implementing security solutions complying with the security level required by different sets of data, access and control monitor processes inclusively etc. All of these can be considered a truism and are basics issues for anyone working more or less close to the information security domain. Despite of this, the practice revealed alarming gaps in the practice. Data protection is in fact composed by a set of policies, processes and solutions interrelated, each representing a link in the security chain and the level of security is always provided by the most fable link of the chain. We all recognize this, but in the practice we forget it so quick. To sustain the above statement, the quality's analysis of the following policies can provide indubitable arguments:

- Organizations seem to be sometime too generous with access rights aiming at facilitating employees' productivity, forgetting the potential negative consequences induced by this fact. Employees have access to more applications and data that really they need to accomplish their jobs. This is the result of an inadequate access right policy. If we corroborate this with breaches in logical access procedures' observance (the well known bad habit to write the password in accessible places or even to divulge it to a colleague for work purposes) we'll obtain two frequent but critical security breaches.

- To protect data integrity and confidentiality companies need a clear, adequate and strictly applied data classification system and corroborated with it a data retention policy. All data and information have to be classified so that the security level ensured for each item to be aligned with the data sensitivity. It is also imperious necessary to be issued and periodically reviewed the data classification system and the retention policy. The policy of data retention, destroy and encryption is linked with data classification. The absence of the retention policy and an inadequate or not periodically updated data classification expose the company to unnecessary and unforeseeable risks. The policy itself must be accompanied by clear processes to make it operational. Data classification and retention policy are ensuring also cost savings related to data security solutions (which varies based of the sensitivity of data) and longest that necessary data retention. Conformity risks are also related to the data privacy and retention issues. International (UE's documents for example but not only) and national lows and rules are regulating data privacy and retention aspects.
  A special notice is imposed for new information technologies as for example cloud computing. Companies are moving data and/or processes in public cloud. Next to the recognized benefits this shifting is bringing potential conformity risks related to special regulated data privacy and retention. In the public cloud data is changing the storage location and this implies conformity risks related to different jurisdictions' regulations. Even encryption and decryption are also subject of specific regulations in different countries.
  The access rights will be also granted based on the data sensitivity; the most sensitive data will be subject of stronger security controls and access granted to a limited number of users.
- More and more organizations are aware that the user's ID and password are not effective control access elements. In this respect, the companies are shifting to another authentication method: 2-factor authentication implying next to the user ID and password another component as for example a smart card in the user's possession or a biometric characteristic (as for example the fingerprint, face and/or voice recognition etc).

CISOs recognize that they are facing with majors barriers in ensuring information security. These are (Deloitte, 2010):

- lack of sufficient budget;
- increasing sophistication of threats;
- emerging technologies;
- lack of support from lines of business.

In the specialists opinion, there are some solutions that can improve the information security:

- implementing a centralized security function and establishing a centralized security model;
- strong security information and event management;
- setting an adequate information security metrics aligned to business value and their report on scheduled basis;
- convergence between information and technology risk functions;
- increasing the information security function visibility inside the company by demonstrating the direct relationship between information risk and business's success.

Aiming at addressing in the most appropriate manner the present dynamic thereat landscape, CISOs are embracing sound risk practices designated to focus the changing IT environments. This new approach known as real-time risk management (RTRM) allows CISOs to obtain up-to-date security information reflecting the current status of their IT environments and detect malicious activity in time so that to minimize their impact.

Another new approach called real-time threat management (RTTM) is dedicated to provide detailed information on the networks' behavior aiming at identifying anomalies in traffic patterns and data flows. In case of suspicious network activities, RTTM can automatically take preventive actions such as quarantining suspicious files and "cleaning" infected endpoints.

4. Aligning and integrating IT security with business strategy

Aligning and integrating IT security with business strategy is an important desideratum of nowadays. IT security governance mechanisms have to be aligned to the operational needs and shape the business model. Business IT-based implies compliance with security standards. In this respect, the

security policy has to be aligned to the security standards and/or regulations specific for the existing business IT solutions.

Global business is arising more complex IT risk compliance issues. The security policy has to meet multijurisdictional compliance requirements and implies more complex and expensive solutions. Remaining aligned to those compliance requirements is cost demanding and time consuming. A good and sustainable security policy has to be flexible, build on security components that are adjustable and easy to assimilate emergent technologies in cost-effective solutions. The business itself is changing and its IT support infrastructure has to follow the business change by redesigning IT processes and automated controls. In this respect, the security policy will be adjusted aiming at facing new emerging risks. The quality of IT controls is important because they ensure: the costs' control, competitiveness, information assets' protection and the compliance with regulations.

To deliver improved information security and risk management, senior management in charge with governance must ensure that the IT strategy is aligned with the business strategy based on thereat landscape and security priorities (Stanciu, 2012).

Changes in IT environment, as response to the business continuous remodeling, determine mutations in risk landscape. As a result, the IT threats must be reanalyzed periodically. The process aims at identifying new emerging threats and their impact over the IT environment and per consequence over the business. Measuring, on regular bases, the business risk profile, in e-cyber domain in particular, is essential in processes' trust security levels. In this respect, new controls or adjusting existing controls decisions are taken aiming at mitigating the IT risk exposure. Implementing the right controls in the right places is the key in optimizing the security controls in terms of cost-effectiveness.

5. Conclusions

The economic and financial world is so dynamic and complex today. One of the factors is represented by the huge IT development and its business implementations. Traditional business models are gradually replaced by models essentially based on software. Physical boundaries are disappearing as more business data is transmitted over the internet and even business models are rethought, data and operations being shifted on cloud platforms. The "volatility" of the physical boundaries is determined in a significant manner by the widespread adoption of mobile devices. In this context, a well governed risk acceptance process is essential for the decision making process and business management process as a whole. IT strategy is considered in many companies as one of the most strategic concern for business. In this respect a proactive attitude in IT risk management is advisable. The information security has to be integrated into decision making process. The level of security investments depend on the management risk appetite and understanding of information technology risks.

As information security function matures is recognized as a strategic necessity for the business. There is still room for imbedding IT risk in decision making process and aligning security goals to those of the business. Till line business managers will not be involved in security issues related to the supporting information systems used in their responsibility area, the merger between information security and business will continue to remain a just desiderate.

Companies recognize their concern for IT risks and they start to do more in this area, but the reactive approach remains the most popular: companies start to invest more in IT after they faced with severe security incidents. In opposite a proactive approach is the one offering advantages protecting IT assets and company's business and reputation.

Strong IT security key stays in a centralized, global solution build at the company's level and accepting the fact that there is no unique solution that may be put in place. The company has to think its own policy and decide on the security solutions selecting the ones that fit to the business.

Investing in IT security is a benefic decision. The investment implies not only hardware infrastructure and software but people also. An important return of investment is provided in case of the investment in employees' education and permanent training.

Adopting an integrated view of IT risk represents a milestone of the risk management process. All senior managers must be aware of the importance of an adequate IT risk management and allocate the needed resources for the IT function and security solutions in order to keep IT risk at an acceptable level.
As IT function matures and its contribution to the business success will be more visible and accepted it will acquire the worthy role inside the company. This is also a matter of organizational culture and needs a new understanding, from the top management part, of the fact that IT became an integral component of

the business and management process too and that IT risks have to be taken into consideration in decision making process.

No matter how complex the implemented security solutions can be a potential risk over the IT systems still exists. Everybody is aware that there is no perfect security solution. The security policies and solutions must be tailored on the specificity of the business model and environment. All the security solutions must be generated by the holistic security policy and reflect a truth so frequently and sometimes dramatically emphasized by the reality: the security level is given by the weakest link.

6. References

- Deloitte (2010) *2010 Financial Services Global Security Study. The Faceless Threat*. Available at http://www.deloitte.com/view/en_GX/global/industries/financial-services/2ac300d256409210VgnVCM100000 ba42f00aRCRD.htm
- Ernst & Young (2011) *Into the cloud, out of the fog: Ernest & Young 2011 Global Information Security Survey*. Available at http://www.ey.com/GL/en/Services/Advisory/2011-Global-Information-Security-Survey---Into-the-cloud--out-of-the-fog
- Good Technology (2011) *Good Technology State of BYOD Report*. Available at http://www.itworld.com/mobile-wireless/247888/good-technology-state-byod-report
- IT Governance Institute (2009) *Enterprise risk: Identify, Govern and Manage IT Risk. The risk IT framework*. Available at www.itge.org
- Kaspersky (2011) *Global IT security risk.* Kaspersky Lab ZAO. Available at http://www.kaspersky.com/downloads/pdf/kaspersky_global_it_security_risks_survey.pdf
- KPMG (2011) *The e-crime report 2011. Managing risk in a changing business and technology environment*. E-crime congress. Available at http://www.e-crimecongress.org/global/documents /e-Crime%20Report.pdf.
- Oltsik, J. (2012) *New demands for Real-time risk management*. Enterprise Strategy Group. http://www.trendmicro.com/cloud-content/us/pdfs/business/white-papers/wp_esg-real-time-risk-mgt.pdf
- Shuester, E.; Kalamarides, A. (2012) Thriving in a BYOD world. IntelliClear Inc. Available at http://www.intelliclear.com/pdfs/INCLRBYOD032312.pdf
- Stanciu, V; Bran, F. (2010) *IT Risk*. Annals of Ovidiu University, Economic Sciences Series, volume X (2010), issue no.1, pag.802-806.
- Stanciu, V; Bran, F. (2012) *Emerging IT risks – Are we prepared to proper react?* International Scintific Conference Economical and ecological dimention of the development in the third millenium. Academy of Economic Studies, Faculty of Agrifood and Environmental Economics, 10th of April 2012.
- TrendMicro (2011a) *Virtualization and cloud computing. Security threats to evolving data centers.* Trend Micro Inc. Available at http://www.trendmicro.com/cloud-content/us/pdfs/business/reports/rpt_security-threats-to-evolving-data-centers.pdf
- TrendMicro (2011b) *Virtualization and cloud computing. Security best practice.* Trend Micro Inc. Available at http://www.trendmicro.com/cloud-content/us/pdfs/about/rpt_virtualization-cloud-computing-security-best-practices.pdf
- Underwood, K.; Slemp, C. (2012) *The current state of IT security and Privacy Policies and Practices*. Protiviti Inc. Available at http://www.protiviti.com/en-US/Documents/Surveys/IT-Security-Privacy-Survey-Protiviti.pdf

SIMULTANEOUS EQUATIONS MODEL USED IN STUDYING THE ANALYSIS ON HOUSEHOLD FINAL CONSUMPTION

ZAMFIR Cristina Gabriela
Ph. D. Assistant professor, Faculty of Economics and Business Administration/Department Economics, Dunarea de Jos University, Galati, Romania, cristinagabrielazamfir@yahoo.com

LUPAŞC Adrian
Ph. D. Assistant professor, Faculty of Economics and Business Administration/Department Economics, Dunarea de Jos University, Galati, Romania, alupasc@ugal.ro

LUPAŞC Ioana
Ph. D. Assistant professor, Faculty of Economics and Business Administration/Department Economics, Dunarea de Jos University, Galati, Romania, ioanalupasc22@yahoo.com

Abstract: *This paper presents an analysis on certain problems concerning household final consumption. In this paper, starting with a series of hypothesis and economic relations, we will define a simultaneous equations model, which will analyze the household final consumption. This model will consist of equations that follow the sociological behavior of the population, other will present the financial behavior and last but not least, identity equations (equilibrium equations). The model was developed centering the household final consumption and it started with an equation defining the consumption behavior, leading to the analysis of every factor, using other economic and econometric relations. The system's equation parameters were estimated using quarterly data series, that were adjusted using TRAMO/SEATS, method based on ARIMA models. The estimation for equation parameters was achieved using procedures implemented in EViews.*

Key words: simultaneous equations model, final consumption of the population, consumer behavior

JEL classification: C 10, C 30, C 50, E 21

1. Introduction

In a market economy, household consumption study is central to social and economic analysis, due to many factors that adversely affect its evolution. Among the factors that can affect the consumption function are economic factors that are easily quantifiable, but also psychological and social factors. The analysis of consumption functions is of great importance, because every change made to function parameters may have important implications for economic policy.

Over time, the consumption analysis has been modified by the introduction of other elements such as sociological and psychological factors, or elements concerning the income and time dynamics, thus leading to the development of a new theory for tracking consumption. In this new theory, the consumption analysis has changed from treating the consumption as an independent actor in the economy, to rather studying its behavior taking into account the changes taking place in the markets with which it interacts in one form or another (Karl, Fair, 2006).

Orientation to consumer needs requires the reference point to be a thorough knowledge of these requirements, monitoring them systematically and even anticipating them, all of which must be done scientifically, by designing and implementing appropriate research tools. All of this led to the development of disciplines such as: economic cybernetics, econometrics and marketing. Regardless of the discipline that deals with consumption, one should consider the following issues:

- the dynamics of consumption behavior, because every individual has a certain behavior that changes depending on certain economic conditions, such as: disposable income, state tax policy, consumer prices for goods and services and other factors;
- behavior is influenced by the interactions one has with social and economic environment. How a consumer manifests on the market may lead to changes and on other participants in the sale;
- consumers differ among themselves and they act rationally.

Consumer behavior is the result of "inputs" and of psychological processes taking place in the human mind. In order to address this, a large amount of information concerning the consumption process is required, but also, the factors that can influence behavior must be taken into consideration.

Consumption and hence consumer behavior have enormous influence on the economy through the fact that many economic systems, and also the level of productivity of a business and marketing, are determined by the decision of a consumer in choosing a product. The choice a consumer makes is influenced by several factors, including changes in income level, changes in prices of other goods, buyers forecast on price trends and income, consumption tradition, buyer's preferences. Summarizing all economic theories and models that analyze consumption, one can say that consumption is influenced by economic factors such as: current income, wealth (measured by individual savings), expected income, interest rate, but also social and psychological factors, such as: the desire to achieve other social levels or to maintain a certain position in the same social group (Oprescu, Imperato, Andrei, 2004).

Nowadays, when the phenomena that mark the economic environment are influenced by globalization, interdependence and interaction, governments and economic analysts should look more closely upon the consumer and his direct needs. The changes that took place during recent years, through their influence on population income, have had an impact on living standards and hence consumption. For this reason, the individual should be regarded as a central point in the economy, as a basic element of a chain, or rather a new food chain because he is not just a taxpayer but, influenced by many other factors, and after payment of taxes, he becomes a consumer. This is primarily due to the remaining budget that he will devote to the acquisition of goods and services required. In other words, the individual will be transformed from a simple taxpayer into a consumer, whose decisions regarding purchased goods are extremely important.

In Romania, the continuing decrease of purchasing power, disposable income, but also the increase in inflation, prices and tariffs, have led households to reduce consumption to a minimum or to try to keep, as much as possible, previous levels of consumption, situation that has acted against the economy. Although these phenomena have a great influence on consumption, they manifest differently for each person or household. Some decide that to save, thus decreasing consumption, others decide to consume more, without saving anything. Those who save now will be able to later use the savings to cover consumption and all other necessities.

The aim of this paper is to present a model of simultaneous equations system to monitor the relation between consumption and other economic factors that directly or indirectly influence it.

2. Model definition

Simultaneous equations systems are used in modeling economic phenomena, given that it is impossible to formally describe a phenomenon only with an econometric model consisting of only one unifactorial or multifactorial, linear or nonlinear equation. The only condition that is required for simultaneous equation models, as for the other equations, is that the estimation of equation parameters to be made using data sets that were obtained either from statistical surveys or from data provided by specialized institutions that deal with studying such phenomena.

The major problem of simultaneous equation systems is that they have a large volume, that is, consist of a large number of equations. In turn, each equation assumes a linear or nonlinear relation between one or more economic factors. The choice of instrumental variables is also difficult because of the necessity to consider all possible relations that may exist between them, so that the model size increases, and also increases the difficulty in determining the parameters (Davidson, Mackinnon, 2004).

The development of the simultaneous equations model begins with determining the desired objective, and continues with specifying the set of starting assumptions. The equation system that we will present further on was defined around household final consumption, taking into account the economic relations between consumption and other economic factors.

Model hypothesis

In order to analyze household final consumption a series of hypothesis were considered:

Hypothesis 1. Household final consumption is an important component of the gross domestic product and it is determined by available total income, adjusted with the inflation rate.

Hypothesis 2. There is a balance relation between aggregate demand and gross domestic product.

Hypothesis 3. There is a balance relation between primary budget deficit and financial wealth.

Hypothesis 4. The statistical data series used in estimating model parameters are obtained from expert sources such as national or international statistical centers.

Hypothesis 5. The statistical data series are adjusted, that is, seasonal phenomena characteristic to raw data series will be eliminated using additive or multiplicative regression models.

Model equations

Model equations were defined starting with the previous assumptions, but also with the economic relations between economic factors (Dobrescu, 2006). The first equation describes the consumption behavior of the population (C_t). Household consumption depends on various economic factors which will be grouped as follows: total disposable income (Y_t^D) containing, according to the identity relation, the gross domestic product (Y_t) and transfers (Tr_t - income from social benefits, allowances for different type of people or unemployment benefits) from which to subtract taxes that must be paid (Tx_t); financial wealth (A_t) represented by bank deposits, shares on the stock market or investments in mutual funds and pension, plus current inflation (π_t) that adjusts prices and income. Given the economic factors mentioned above, the first regression equation will have the following form:

$$C_t = f_1(Y_t^D, A_t, \pi_t) + \xi_{1t} \quad (1)$$

where f_1 is a linear or nonlinear combination between the exogenous variables of the equation and ξ_{1t} is the residual variable, not correlated with explanatory variables. As stated previously, the identity relation for disposable income has the following form:

$$Y_t^D = Y_t - Tx_t + Tr_t \quad (2)$$

The second equation estimates the role of various factors in determining the taxes demanded by the state. The factors that were taken into account to determine the equation were: GDP, which is adjusted with a tax rate (τ) and autonomous taxes (To_t). The system regression equation will have the following form:

$$Tx_t = \tau Y_t + To_t + \xi_{2t} \quad (3)$$

where ξ_{2t} is the random variable, that has a normal distribution and the tax rate is the parameter of the exogenous variable GDP, while the autonomous taxes will be determined as a parameter for the regression equation.

The third equation of system is the relation between inflation and unemployment, equation identified as the Phillips' curve. Phillips' curve is a relation between the growth rate of monetary wage and the employment level:

$$\pi_t = i_\pi \pi_{t-1} - i_u u_t + \xi_{3t} \quad (4)$$

where ξ_{3t} is the random variable and the employment level is represented by the unemployment rate (u_t), while the wage growth is given by the inflation rate.

The fourth regression equation, part of the equation system is an equation of financial behavior. The regression equation shows the relation between investment (I_t) and income, which is updated in concordance with the inflation rate:

$$I_t = a_0 + a_1 r_t + a_2 Y_t + a_3 A_t + a_4 \pi_t + \xi_{4t} \quad (5)$$

The last two regression equations of the system are those related to the international trade. The first equation shows the relation between exports (X_t) and the real exchange rate (R_t) and the second shows the relation between imports (Z_t), the real exchange rate and GDP.

$$X_t = c + x_R R_t + \xi_{5t} \quad (6)$$

$$Z_t = z + z_Y Y_t - z_R R_t + \xi_{6t} \quad (7)$$

where ξ_{5t} and ξ_{6t} are errors and they are appear in the model as the sum of all unknown influences.

Summarizing all seven equations, the result will be a system with the following form:

$$\begin{cases} C_t = c + c_V Y_t^D + c_A A_t + c_\pi \pi_t + \xi_{1t} \\ Y_t^D = Y_t - Tx_t + Tr_t \\ Tx_t = \tau Y_t + To_t + \xi_{2t} \\ \pi_t = i_\pi \pi_{t-1} - i_u u_t + \xi_{3t} \\ I_t = a_0 + a_1 r_t + a_2 Y_t + a_3 A_t + a_4 \pi_t + \xi_t \\ X_t = c + x_R R_t + \xi_{5t} \\ Z_t = z + z_Y Y_t - z_R R_t + \xi_{6t} \end{cases} \quad (M1)$$

in addition to these, there will be a series of identity equations to strengthen the model (M1).

$$\begin{cases} D_t = C_t + I_t + G_t + Nx_t \\ DBP_t = G_t + Tr_t - Tx_t \\ Nx_t = X_t - Z_t \\ R_t = \dfrac{Pf_t e_t}{IPC_t} \\ D_t = Y_t \\ A_t = M_t + B_t \\ DBP_t = A_t \end{cases}$$

The first equation of the seven, expresses the calculation of aggregate demand. Taking into account the second hypothesis stated in the first part of the paper, the aggregate demand equals supply, represented by GDP, as shown in equation five of the system. Aggregate demand is calculated as the sum of consumption, investment, government expenses (G_t) and net export (Nx_t) and expresses the GDP in terms of expenses.

The second equation shows the primary budget deficit (DBP_t) and given the third hypothesis, it is assumed that there is an equality between wealth and gross primary deficit. Considering the stated hypothesis, the wealth calculated as the sum of money mass (M_t) and bonds' issue (B_t) will be equal to the primary budget deficit, which is calculated as the difference between the total public expenditure, consisting of those for purchasing goods and services (G_t) plus transfer payments (Tr_t) and the amount of tax revenue (Tx_t).

The third identity equation expresses the net export equation (Nx_t), calculated as the difference between exports and imports. Because both import and export are two variables that, as stated in the regression equation, depend on the real exchange rate, this will be calculated as the ratio between foreign consumer price index (Pf_t) multiplied by the exchange rate (e_t) and domestic consumer price index (IPC_t).

The model aims mainly to measure the influence of certain economic factors on consumption, but also to pursue the relations between these factors. This system of simultaneous equations consists of fourteen different equations: a behavioral equation (equation 1), a financial behavior equation (equation 5), six equations of economic identity, an equation showing the economic balance (IS balance) and a relation showing Phillips' curve. The model structure includes ten endogenous and six exogenous variables. The endogenous variables of the model are represented by consumption, disposable income, GDP, taxes, wealth, inflation at the current time, investments, net export, exports and imports, and the exogenous variables are represented by transfers, government expenses, inflation at a later time, unemployment, real interest rate and exchange rate.

Because to the model (M1) were added a series of economic identity equations, the reduced form of the model can be reformulated as follows:

$$\begin{cases} C_t = c + c_V Y_t^D + c_A A_t + c_\pi \pi_t + \xi_{1t} \\ Y_t^D = Y_t - Tx_t + Tr_t \\ Y_t = C_t + I_t + G_t + Nx_t \\ Tx_t = \tau Y_t + To_t + \xi_{2t} \\ A_t = G_t + Tr_t - Tx_t \\ \pi_t = i_\pi \pi_{t-1} - i_u u_t + \xi_{3t} \\ I_t = a_0 + a_1 r_t + a_2 Y_t + a_3 A_t + a_4 \pi_t + \xi_{4t} \\ Nx_t = X_t - Z_t \\ X_t = c + x_R R_t + \xi_{5t} \\ Z_t = z + z_Y Y_t - z_R R_t + \xi_{6t} \end{cases} \quad \text{(M2)}$$

The equation parameters of the system were determined using the two-phase method of least squares. This method was chosen because, after evaluating the model, it was found that the model was over-identified, containing six over-identified equations and four economic identity equations. The first phase of the two-phase method of least squares is to define the model in its reduced form. The reduced form of the model requires that all endogenous variables of the model to be expressed only through its exogenous variables, thus obtaining the maximum likelihood estimators. Using the ten endogenous and six exogenous variables the reduced form model was obtained having the following form:

$$\begin{cases} C_t = \alpha_1 Tr_t + \alpha_2 G_t + \alpha_3 \pi_{t-1} + \alpha_4 u_t + \alpha_5 r_t + \alpha_6 R_t + \alpha_7 \\ Y_t^D = \beta_1 Tr_t + \beta_2 G_t + \beta_3 \pi_{t-1} + \beta_4 u_t + \beta_5 r_t + \beta_6 R_t + \beta_7 \\ A_t = \gamma_1 Tr_t + \gamma_2 G_t + \gamma_3 \pi_{t-1} + \gamma_4 u_t + \gamma_5 r_t + \gamma_6 R_t + \gamma_7 \\ \pi_t = \varphi_1 Tr_t + \varphi_2 G_t + \varphi_3 \pi_{t-1} + \varphi_4 u_t + \varphi_5 r_t + \varphi_6 R_t + \varphi_7 \\ Y_t = \delta_1 Tr_t + \delta_2 G_t + \delta_3 \pi_{t-1} + \delta_4 u_t + \delta_5 r_t + \delta_6 R_t + \delta_7 \\ Tx_t = \psi_1 Tr_t + \psi_2 G_t + \psi_3 \pi_{t-1} + \psi_4 u_t + \psi_5 r_t + \psi_6 R_t + \psi_7 \\ I_t = \mu_1 Tr_t + \mu_2 G_t + \mu_3 \pi_{t-1} + \mu_4 u_t + \mu_5 r_t + \mu_6 R_t + \mu_7 \\ Nx_t = \varpi_1 Tr_t + \varpi_2 G_t + \varpi_3 \pi_{t-1} + \varpi_4 u_t + \varpi_5 r_t + \varpi_6 R_t + \varpi_7 \\ X_t = \nu_1 Tr_t + \nu_2 G_t + \nu_3 \pi_{t-1} + \nu_4 u_t + \nu_5 r_t + \nu_6 R_t + \nu_7 \\ Z_t = \omega_1 Tr_t + \omega_2 G_t + \omega_3 \pi_{t-1} + \omega_4 u_t + \omega_5 r_t + \omega_6 R_t + \omega_7 \end{cases} \quad \text{(FR)}$$

where φ_1, φ_2, φ_5, φ_6, φ_7 were proven to be null, the same results being obtained for the υ_1, υ_2, υ_3, υ_4, υ_5 parameters.

Using the equations from the model in its reduced form, we will identify the parameter coefficients of the regression equations for the endogenous variables: GDP, wealth and disposable income (Y_t, A_t and Y_t^D), taking into account the six exogenous variables. After identifying the regression equations for the three endogenous variables, their theoretical values will be calculated. For the evaluation of regression equation coefficients, Romania's quarterly data for 2005-2009 was used. The data sets were adjusted using TRAMO / SEATS models, which are based on ARIMA models. The values obtained for the coefficients and the significance tests for each equation are presented in table 1.

Table 1: The estimated values for the regression equation coefficients for GDP, wealth and disposable income

	Coefficient	**t-Statistic**	**Determination coefficient (R^2)**
β_1	-0,02	0,36	0,458
β_2	-0,08	-0,70	
β_3	84,35	0,27	
β_4	38,98	0,29	
β_5	-118,42	-1,04	
β_6	1189,95	1,99	
β_7	932,53	7,98	
γ_1	0,75	8,33	0,99
γ_2	0,24	2,44	
γ_3	0,05	2,25	
γ_4	0,10	0,98	
γ_5	0,05	0,79	
γ_6	-0,79	-4,02	
γ_7	1,14	6,40	
δ_1	2,56	0,93	0,50
δ_2	-2,49	-0,96	
δ_3	-0,13	-0,17	
δ_4	-1,77	-0,58	
δ_5	2,48	1,21	
δ_6	-3,55	-0,58	
δ_7	0,87	0,89	

The R^2 test, called also the determination coefficient, is intended to demonstrate the likelihood of the model and shows the variation of the random variable. In order to specify that the three equations above are plausible, the determination coefficient must be greater than or equal to 0.50. As can be seen in table 1, the values of this coefficient for each of the three equations, are near the value of 0, 50, so we can say that the regression equations determined above are plausible.

In determining the correlation between exogenous variables and the random ones, the Durbin-Watson test was used, for a significance level of 5%. All three equations verify this error autocorrelation hypothesis, so that we can say that there is no link between the two variables of the equations and the estimated values for the parameters are correctly determined.

The regression equations determined for the three exogenous variables, will continue the second phase of the two phase method of least squares. Using the new values determined according to the identified equations for the three exogenous variables, the paper will continue with the calculation of the regression equation for household final consumption, as shown in equation (1).

Before determining the coefficient values, dependence was found between the exogenous variables of the consumption equation, phenomenon which was eliminated by differentiating the economic factors, resulting in a model with the following form:

$$C_t = c + c_V (Y_t^D - Y_{t-1}^D) + c_A (A_t - A_{t-1}) + c_\pi (\pi_t - \pi_{t-1}) + \xi_t \quad (8)$$

Using the regression equation (8), which eliminated the multicolinearity phenomenon, the coefficient values were determined and the following values were obtained:

Table 2: The estimated values for the regression equation coefficients of household final consumption

	Coefficient	t-Statistic	Determination coefficient (R^2)
C	-9272,28	4,59	0,359
c_V	5,67	-1,94	
c_A	0,46	-0,81	
c_π	255,61	-0,55	

The homoscedasticitity hypothesis, that is verifying that the residual variable has an independent distribution, was conducted using the White test, by checking the calculated value of Fisher-Snedecor test and the tabulated value for the twelve remaining observations.

The error autocorrelation hypothesis, that is verifying that errors aren't correlated (correlated errors mean that the coefficients are not properly determined), was checked with the Durbin Watson test. The value obtained for the twelve observations was equal to 2.13, value which is within the interval that states the errors are independent, so the error can not perpetuate on to the values of variables.

Although the correlation coefficient is low, we can say that there is a link between exogenous variables of the model and the calculated variable. The model is plausible and can be accepted to track the relation between household final consumption and exogenous variables: wealth, disposable income and inflation.

3. Conclusions

Starting with a few assumptions, a simultaneous equations model was created. This model can analyze the household final consumption taking into account a limited number of macroeconomic factors. Transition from a structural form to a reduced form was done taking into account all the exogenous variables of the model. Once the reduced form of the system was determined, we estimated values for each variable that appeared in the consumption equation. Each of the regression equations presented in this paper were checked in terms of plausibility, that is, all hypothesis required for the type of model were checked.

The household final consumption was determined using exogenous variables values, which were calculated after reformulating the regression equations, taking into account the reduced form of the model. This model became convergent after several iterations and has the following form:

$$C_t = -9272{,}28 + 5{,}67 * Y_t^D + 0{,}46 * A_t + 255{,}61 * \pi_t \quad (9)$$

In conclusion, household final consumption is affected by any change that takes place throughout the economy. Any information about price or unemployment increase, wage decrease or even global economic situation (economic crisis) may lead to a decrease in consumption, which is normal. Analyzing data sets from the National Center for Statistics, it is noticed that in 2001, household final consumption represented 78.15% of total gross domestic product (GDP) and after a year, it decreased by about 1%, reaching to 76.9% of GDP, while the total household monthly income increased by 26.20%.

4. References

- Alexandrescu, P. (2007) *Introducere în statistică socială*, Editura Paralela 45, Bucureşti
- Baltagi, B.H.(2008) *Econometrics 4th edition*, Springer
- Davidson, R; James G. M. (2004) *Econometric theory and methods*, Oxford Press

- Dobrescu, E. (2006) *Macromodels of the romanian market economy*, Editura Economică, Bucureşti
- Gujarati, D. N. (2004) *Economics–basic econometrics*, McGraw-Hill Companies
- Iacob, A. I. (2004) *Econometria consumului populaţiei*, Editura ASE, Bucureşti
- Ivan, I.; Vişoiu A. (2005) *Baza de modele economice*, Editura ASE, Bucureşti
- Karl, C.; Ray C.F. (2006) *Principles of macroeconomics*, ediţia 8, Prentice Hall
- Kirchgässner, G.; Wolters J. (2007) *Introduction to modern time series analysis*, Springer
- Mankiw, G (2007) *Principles of macroeconomics*, South-Western CENGAGE Learning
- Oprescu, G.; Imperato A.; Andrei A. (2004) *Cibernetica consumatorului. Abordare statică şi dinamică*, Editura ASE, Bucureşti
- *** *Anuarul Statistic al României,* Editura Institutului Naţional de Statistică, ediţiile 2005, 2006, 2007, 2008, 2009

SOA CONSIDERATIONS FOR AN ELECTRIC POWER INDUSTRY ENTERPRISE

ZOTA Răzvan Daniel
Professor, Ph.D., Faculty of Cybernetics, Statistics and Economic Informatics/Department of Economic Informatics and Cybernetics, Bucharest Academy of Economic Studies, Bucharest, Romania, zota@ase.ro

MOLEAVIN Radu Ştefan
Ph.D. student, Faculty of Cybernetics, Statistics and Economic Informatics/Department of Economic Informatics and Cybernetics, Bucharest Academy of Economic Studies, Bucharest, Romania, radustefan.moleavin@enel.com

CONSTANTINESCU Radu
Assistant professor, Ph.D., Faculty of Cybernetics, Statistics and Economic Informatics/Department of Economic Informatics and Cybernetics, Bucharest Academy of Economic Studies, Bucharest, Romania, radu.constantinescu@ie.ase.ro

Abstract: *The electricity networks used today in the electric power industry are still using outdated electromechanical equipments, are unsafe, maintenance is expensive and the time for it cannot be controlled easy. To solve these problems the electricity networks must become more "intelligent" by introducing electronic control devices and self-healing mechanisms for better safety and efficiency. The new equipments and devices must also be interoperable and right integrated into business; this may be achieved by using software oriented architecture (SOA). In this paper we describe a service oriented architecture designed for an enterprise from the electric power industry and some conclusions about it.*

Key words: SOA, Business Processes, Web Services, Electric Power Industry, Enterprise

JEL classification: D80

1. Introduction

For every electrical distribution enterprise, the most important constraints are stability and electricity quality. Stability can be achieving by improving SAIDI (System Average Interruption Duration Index) and SAIFI (System Average Interruption Frequency Index). The quality of electrical power could be described as a set of parameters' values, such as: voltage stability, continuity of service, voltage magnitude variation and others.

A prompt diagnosis of the problems that occurs in the system distribution of electricity leads to their prompt correction. The electricity networks are complex and the volume of information comprised is huge. As events occur on electricity networks, having the right information about it is vital in monitoring and controlling them. To manage the information in the electricity networks we need to use information systems like Supervisory Control and Data Acquisition (SCADA), Distribution Management Systems (DMS) and Energy Management System (EMS). Monitoring the electricity networks can lead to reliable measures for the prevention of disturbances. In order to control and monitor electricity networks these systems have to receive data from field devices of the electricity networks.This means automation of the field devices which include fault detection, localization, isolation and load restoration, which will detect a fault, localize it to a segment of feeder, open switches around the fault and restore unfaulted sources via the substation and alternative sources as available.

Equipments for automation of the field devices can be located in various parts of the distributions networks and thus can be monitored. To have data from the information systems, the field devices need to be integrated and to be interoperable each other. Since the existing distribution system is highly complex and widely distributed, it is needed to develop a service oriented open system made up of a variety of power system services operating on different platforms. In this way, more and more extensive data and applications can be shared easily and flexibly.

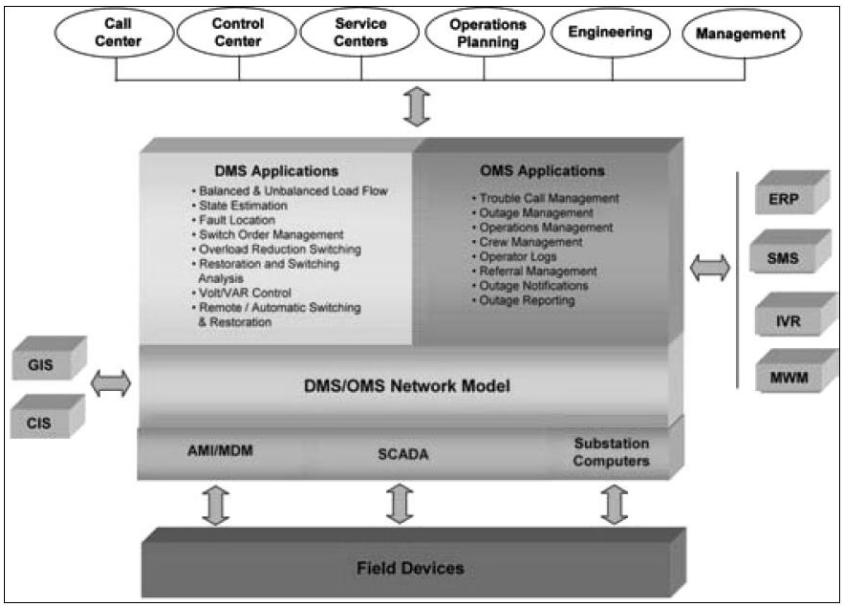
Furthermore, the SOA model is a good solution for distribution networks stability operations because it provides a powerful set of features based on open standards like XML, SOAP, WSDL and UDDI and supports integration and interoperability between services on different platforms in a reliable, scalable and adaptable manner.

2. A general architecture for a distribution enterprise

A Distribution Operations Center is composed of following information systems: a Distribution Management System (DMS), an Outage Management System (OMS) and a Supervisory Control and Data Acquisition (SCADA). The functionality of a DMS is a new fact which implies that the applications of such a system are being used in the outage management processes. In the same time, a DMS system is extending the efficiency of the planned work management and for the common electricity operations. A DMS is typically associated with a real-time status inventory and analogue points from the distribution system. Also, it acts as a supervisory command control deck for the distribution breakers, switches and reclosers, switched capacitor banks, voltage regulators, and load tap changers (LTCs).

SCADA (Supervisory Control and Data Acquisition) systems are usually seen as a method for real-time monitoring and control of the electrical power systems and in particular for the generation and transmission systems (Taylot, 2009). Terminal like RTUs (Remote Terminal Units) are used for the collection of telemetry data (both analogue and status) from field devices. They are also used to pass control commands to general field devices. Generally speaking, an OMS (Outage Management System) has generic functions like trouble-call handling, outage analysis and prediction, reliability reporting and crew management. The OMS systems have been also integrated with other operational systems such as Geographic Information Systems (GIS), Customer Information Systems (CIS), Work Management Systems (WMS), Mobile Workforce Management (MWM), SCADA and AMI. The integration of an OMS with these systems had led to good results in better workflow efficiency and enhanced customer service.

Figure 1: Architecture for a fully integrated Distribution Operations Center



Source: (Taylot, 2009)

DMS and OMS applications are using a usual network model. If the OMS applications are used primarily in outage response situations, the DMS applications are typically related to the electricity operation of the network and use electricity data from the integrated DMS/OMS model. Examples, in that case, include line and cable impedances, equipment ratings and customer load characteristics. The DMS/OMS can use data from other IT distribution systems which include SCADA. SCADA systems continue to extend over the distribution substations and onto the feeders, while offering improved data for awareness and control. There have been also provided interfaces between AMI/MDM (Advanced Metering Infrastructure/Meter Data Management) and the OMS. These interfaces have the role of metering pinging, restoration and outage notifications.

The possibility of using other AMI data in the DMS applications (like *interval demand* and *voltage violations*) is taken into consideration. Moreover, some organizations are using more and more substation automation and substation computers integrated in their systems. This fact leads to a raised

access to data in the intelligent electronic devices (IEDs) that are being installed in the distribution system and in substations. As a consequence, these "include more intelligent recloser controls, switch controls, and voltage regulator controls" (Taylot, 2009). The way data is transmitted between the field devices and the integrated operations centre may be different for the organizations from the distribution field, and there also may be different methods within the same company. No matter the approach being used, the output data will better assist in increasing the operational awareness of the whole system.

3. Service oriented architecture

The problems depicted in a normal architecture of an enterprise from the electrical industry can be solved using SOA, which is defined in (Josuttis, 2007) as: „an architectural paradigm for dealing with business processes distributed over a large landscape of existing and new heterogeneous systems that are under the control of different owners".

SOA is essentially a distributed architecture, with systems that include computing platforms, data sources and technologies. A distributed architecture requires integration. Integration software provides the bridge between the legacy systems and SOA, SOA represents a software architecture that is business-oriented and integrates the business tasks as a set of interconnected and reusable services, which communicate with one another.

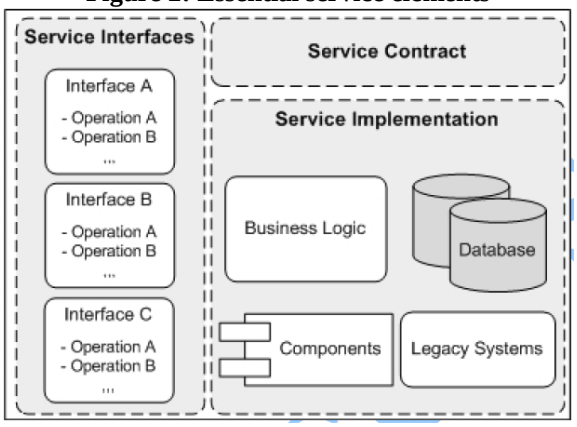
Strategic goals and benefits of SOA are (Erl, 2009):

- *Increased Intrinsic Interoperability*- Interoperability refers to the sharing of data and more interoperable programs exchange easier information.
- *Increased Federation*- A federated IT environment is one which resources and applications are united and have individual autonomy and self-governance.
- *Increased Vendor Diversification Options*. This help enterprise to quickly change business processes to at changes of market.
- *Increased Business and Technology Alignment*
- *Increased ROI* – Services have increased reuse potential that can be realized by allowing them to be repeatedly assembled into different compositions.
- *Increased Organizational Agility* - Enterprise agility on refers to efficiency with which an organization can respond to change.

The basis for SOA is the *service*, described as a function performed by an application. A function is coded only once and then is reused wherever it is needed. As shown in (Qusay 2009), a service is a "self-contained (offers different functionalities related to one business or technical area/sub-area), cohesive (all related functionalities are placed together), black box (consumers know nothing about its internals, and underlying technologies) software component that encapsulates a high level business/technical concept that can cover specific area of whole system".

Technically, a service is composed of three parts, as shown in Figure 2:

Figure 2: Essential service elements



Source: (Qusay, 2009)

• **Contract:** It provides both *formal* and informal specifications of service. Formal specifications use a description language like WSDL to describe information related to technical areas of service such as underlying programming language(s), middleware(s), network protocol(s) and other runtime aspects. Informal specifications are textually presented to provide general information such as the purpose, functionality, constraints, usage of exposed service and expected response time. A contract (Josuttis,

2007) is the complete specification of a service between a specific provider and a specific consumer. From the consumer's point of view, it defines "everything you have to know when using this service," so that (ideally) no doubts remain.

• **Interface:** It provides technical representation of service operations (contain information about public operations, parameters and return types.) that are available to be invoked by clients.

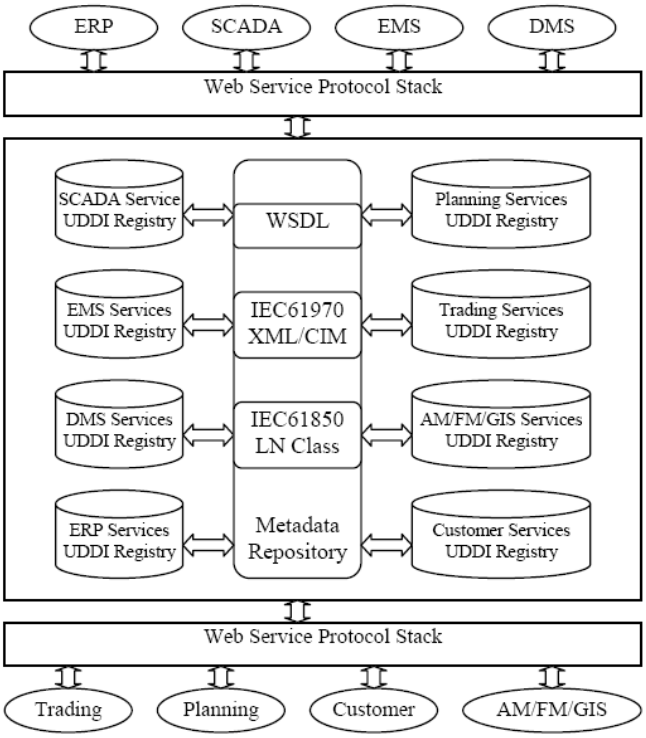
• **Implementation:** It contains logic of service that might be related to accessing data, business logic, etc. This implementation logic could be encapsulated internally within service itself or it may be provided by other external artifacts (e.g. other programs, code libraries, components, legacy systems, etc.).

4. SOA Architecture for Enterprise Distribution

4.1 Introduction

The integration and interoperation within the electric power distribution needs unified common information formats, services and interfaces. The electrical industry makes great efforts to standardize communication protocols and data models. Till now, there are published several standards for power systems and equipments: CCAPI and UCA2.0 by EPRI, IEC61970 and IEC61850 by IEC (Qizhi 2006). SOA, Web services and these information standards make it possible to build an open, flexible and scalable infrastructure for information integration.

Figure 3: Web-services infrastructure for information integration in Enterprise Distribution

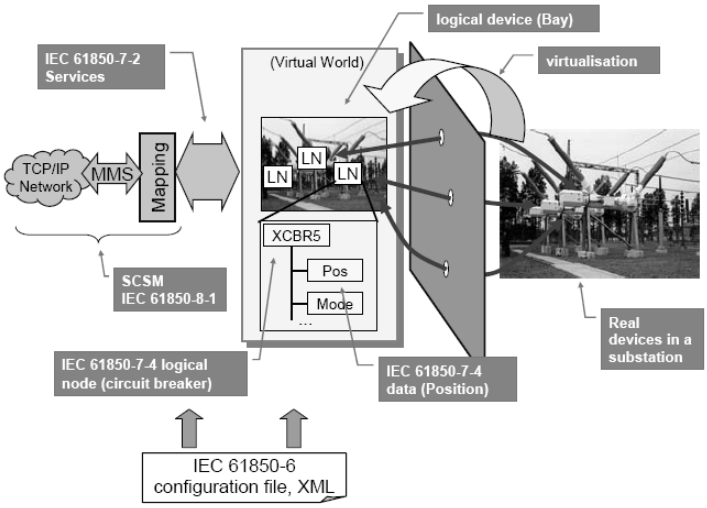


Source: (Qizhi, 2006)

The infrastructure shown in Figure 3 uses the standard components of Web services, i.e. web service protocols stack for communication between different systems, WSDL for Web services description, UDDI for services discovery and integration.

Also, a key element to define the information communication of a device is the Logical Node (LN) (Schwarz, 2004). Each LN provides a list of well organized and named information (e.g. LN "XCBR5" represents the "circuit breaker" number 5 with the data "Pos" (Position) and "Mode"). Services defined in IEC 61850-7-2 allow the exchange of this information.

Figure 4: Logical Node



Source: (Schwarz, 2004)

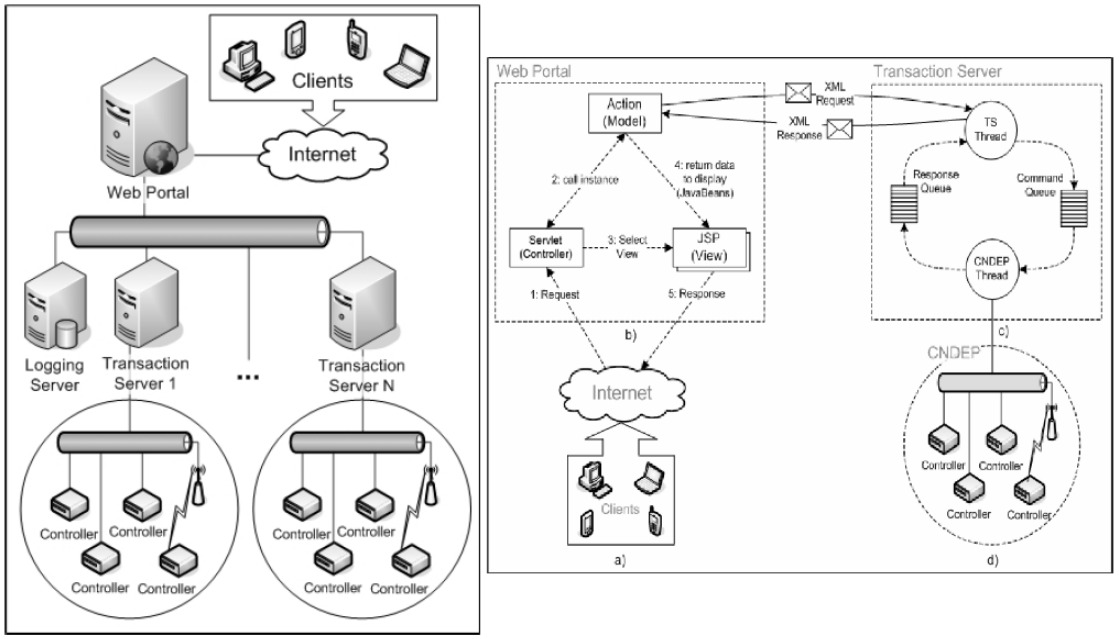
4.2 Distributed Automation Systems

One type from the set of the complex systems which are the foundation for the big enterprise infrastructures today is represented by the Distributed Automation Systems (DAS).

A DAS is composed of many different nodes which need to communicate each other (Kakanakov, 2005). The tasks for these systems require reliable and predictable performance. This approach is now feasible because new technologies are present and on the market today there are a many new automation devices with integrated TCP/IP stack and Web server.

In figure 5 is shown an example DAS Architecture and Functional Model. This example has four tiers:

Figure 5: Example DAS architecture and functional model



Source: (Kakanakov, 2005)

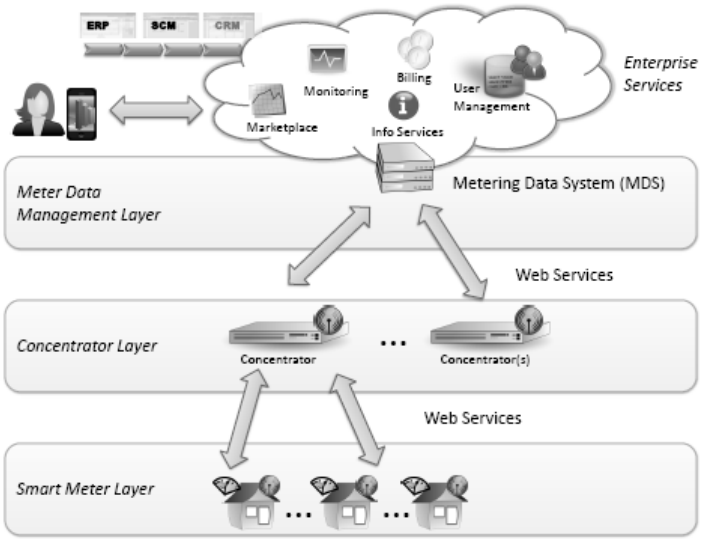
Client tier - A client can be any device with a standard internet browser (PC, Laptop, PDA, Smartphone, etc.). The communication with the Presentation tier is based on exchange of standard HTTP request/response.

Presentation tier - This tier is responsible for handling the requests of clients and forming the view of the responses. After receiving a request, it is analyzed and this is transformed into XML encoded queries and dispatched to the appropriate server from the next tier. Presentation tier is represented by a Web Portal. It is a web server, which acts as point of presence for different services offered by the system.
Services tier - It includes different servers which offers particular services. Services from this tier can be: data logging, plant control and monitoring. Implementation of these services consists of server (on the services tier) and controllers' network (on the data tier). The server receives XML queries and form transactions to the Data tier. This server is called Transaction server. The actual communication with controllers is based on a custom protocol - CNDEP. This protocol is optimized for monitoring and control of automation devices. Transaction server sends commands to controllers and receives results.
Data tier - The Data tier has different expressions for different services on the Services tier. In case of Information services it is a database and in case of Transaction server it is controllers' network. The controllers' network acts as a data producing component. It can implement communication with the controllers in different ways, based on industrial standards or on custom protocols. Each controller in the network can implement monitoring, diagnostic and control tasks.
The Data tier interacts with the environment by means of sensors and actuators, which implements the actual monitoring and control. The role of the controller is to drive the sensors and actuators and to implement the protocol for data extraction.

4.3 Advanced Metering Infrastructure

An Advanced Metering Infrastructure (AMI) using Web Services needs to measure, collect and analyse energy via the usage of meters for electricity (Karnouskos, 2011). As shown in figure 6, the metering data collected from various sources is gathered to strategically positioned concentrators.

Figure 6: Example AMI architecture



Source: (Karnouskos, 2011)

The concentrator represents the interface between many low-speed heterogeneous, usually asynchronous channels and one or more high-speed, usually synchronous, channels. It acts as an interface between the smart meters and the enterprise. It is also responsible for aggregating the meter reading data and submitting it to the metering server. A typical architecture AMI has three-layers as is depicted in Figure 6:

- *The Smart Meter Layer*: here is where the meters are passively tapping in and are measuring the energy consumption of the attached devices.
- *The Concentrator Layer:* the meters pass their measurements to this layer. The reported data is aggregated and submitted to the *metering data system (MDS)*.

- *The Metering Data Management Layer*: here the data and events regarding the infrastructure are collected for long-term data storage, analysis and management. This is typically used by enterprise services in order to empower applications such as billing, fore-casting, and so on.

The implementation of this kind of architecture needs the interconnection of heterogeneous components, namely service-oriented architecture (SOA) or, more specifically, the use of web services. This specific implementation uses web services and Simple Object Access Protocol (SOAP).

5. Conclusions

In order to improve business processes in an enterprise from the electricity distribution industry it is necessary to increase the degree of automation of the electricity networks by using intelligent devices. These devices will provide a large quantity of information recorded from the level field assets to information systems through the internet using web services. The integration can be made by using SOA as an open, flexible, scalable framework with higher cooperation and integration capability that links applications within utilities over the internet. The service-oriented architecture allows flexible composition of functionality: embedded services within smart meters and distribution networks, intelligent devices and higher-level services for business applications. These elements can assure interoperability and compatibility in heterogeneous software environments.

A good SOA implementation can bring benefits to a utility organization in improving agility and in providing the right information to the right stakeholders at the right time in the most scalable and flexible way. By having the right information, the stakeholders may take the right decisions which will lead to business processes improvement. Finally, the implementation will lead to: a reduced cost of interruptions, an increased power quality, reduced electricity production costs and delivery, reduced injuries and loss of life from grid-related events. All these advantages will lead to a major scope: keeping low electricity prices for final consumers.

6. References

- von Dollen, D. (2009) *Report to NIST on the Smart Grid Interoperability Standards Roadmap – Post Comment Period Version.* Retrieved from http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/Report_to_NIST_August10_2.pdf.
- Erl, T. (2009) http://www.soaprinciples.com
- Josuttis, N. M. (2007) SOA in Practice: The Art of Distributed System Design, O'Reilly Media, USA.
- Kakanakov, N.; Shopov, M.; Spasov, G. (2005) *A New Web-based Multi-tier Model for Distributed Automation Systems,* Information Technologies and Control, Vol. 4, no. 2, pp.44-48, 2006, ISSN: 1312-2622, Plovdiv, Bulgaria.
- Karnouskos, S.; da Silva, P. G.; Ilic, D. (2011) *Assessment of High-performance Smart Metering for the Web Service Enabled Smart Grid*. Retrieved from http://www.smarthouse-smartgrid.eu/fileadmin/templateSHSG/docs/publications/2011_ICPE.pdf
- Qizhi, C.; Hamada, G.; Weiming, S. (2006) *Web-services Infrastructure for Information Integration in Power Systems.* Retrieved from http://kpx.or.kr/kems/info/data/Web-Services_Infrastructure_Power_System.pdf
- Qusay, F. H. (2009) *Aspects of SOA: An Entry Point for Starters*. Anale, Seria Informatica, Vol. VII, fasc. 2 -2009. Retrieved from http://arxiv.org/ftp/arxiv/papers/1202/1202.6623.pdf
- Scwarz, K. (2004) *IEC 61850 Communication Networks and systems in substations.* Retrieved from http://www.nettedautomation.com/download/IEC61850-Browsable-Small-Subset-of-Models_2004-05-02.pdf
- Sierra Energy Group (2010) *Smart Grid starts with Smart Design*. Retrieved from http://images.autodesk.com /adsk/files/autodesk_wp_smartgrid_final_en.pdf.
- Taylot, T.; Kazemzadeh, H. (2009) *Integrated SCADA/DMS/OMS: Increasing Distributions Operations Efficiency*, Electric Energy T&D Magazine. Retrieved from http://www05.abb.com/global/scot/scot221.nsf/veritydisplay/722b1bc76a16ae038525759a005f0d65/$file/ee-ma09_abb_02.pdf

THE SCHOOL MANAGER – DECIDING FACTOR IN THE SCHOOL'S ORGANIZATION

MIHĂESCU Diana
Ph.D. Lecturer, "Lucian Blaga" University of Sibiu, diana.mihaescu@ulbsibiuro

Abstract: *In the architecture, functionality and performance of the educational system's organizational management its decisional component occupies a privileged position. No other element has such a peculiar management so evident and with such a great impact in all the activities' areas and its results. The decisional subsystem[1] in educational field it is the component of the educational' management system with the most obvious and specific management, and its quality marks the functionality and the performance of the organization.*

JEL: M12, M19, M20

Given the nature of the tasks involved in the managerial process development of the educational institution the ***educational management functions*** have the following ***characteristics***[2]: are the basic functions of management, developed by all managers regardless the level they act (strategic, tactical, operational); develops differentiated on each level of Maslow's managerial hierarchy pyramid; have different weights in all managerial activities, depending on the managers' level of action (fig.nr.1). However there are other functions of management: belonging to managerial positions, determined by grouping the activities based on more sensitive criteria to differentiation.

Figure 1. Levels of action of school' managers

At the base of the functions of educational management are three basic elements: ideas, things and people. Planning deals with ideas, the organization deals with things and the management and control deals with people. The management of a school organization must struggle for maintaining a fair balance between the functions of management in order to ensure efficient use of resources and market success of the organization.

Systemic treatment of ***management functions***[3] of school organization involve beside taking into account their separate development, as an entity, and the analyzes of the relations between them (fig.nr.2). Separate approach to management functions and of interdependencies between them contributed to highlighting the characteristics of school organization process of management and a certain degree of complexity, including multiple sides that are involved.

[1] Nicolescu O., Verboncu I., *Management*, Ed. Economică, Bucureşti, 1995;
[2] Lazăr I., Vereş V., Mortan M., *Management general*, Cluj-Napoca, Editura Dacia, 2002;
[3] O. Nicolescu, I. Plumb, M. Pricop, I. Vasilescu, I. Verboncu (coord.), *Abordări moderne în managementul şi economia organizaţiei*, vol. 1-4, Ed. Economică, Bucureşti, 2003;

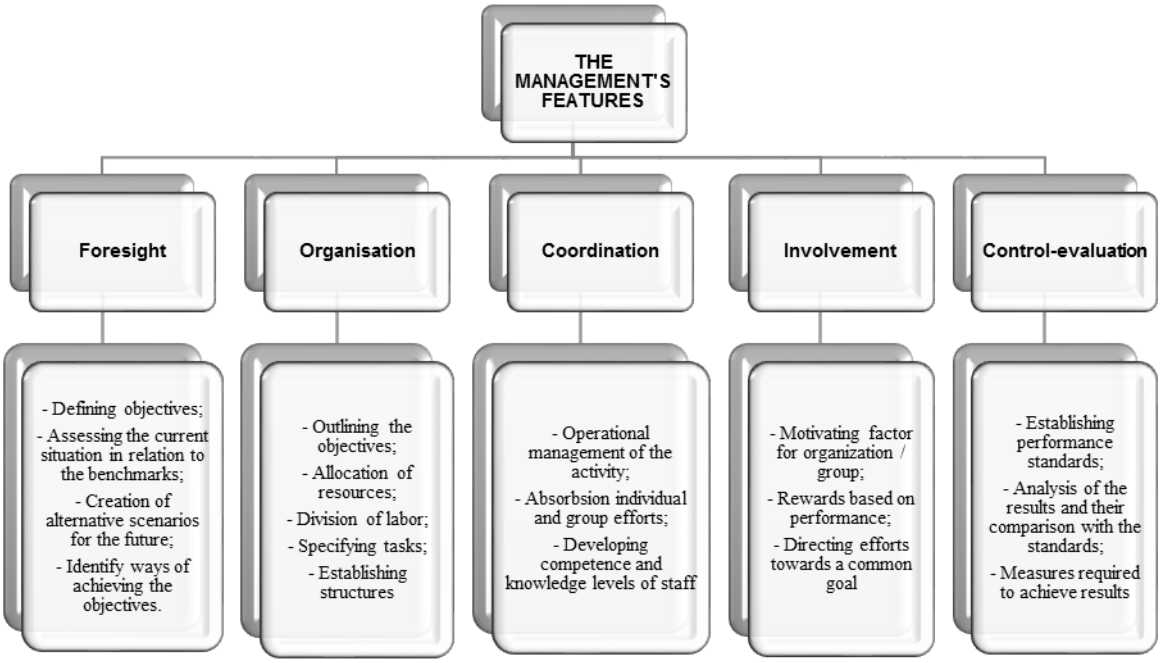


Figure 2: Management features and characteristics of school organization

School organization management features are closely interrelated and form a coherent system of a dynamic nature. They appear and develop along with the development and school organization expressing more or less accentuated its character at different stages of development of the school's organization[4].

Interdependent approach of educational management features is determined by the systemic nature of the school organization over which it is applied so that any deficiencies in management activities is reflected in lower profitability of the organization.

The managerial activity in school organization identifies with the concept of management as a process. It is not a novelty that school manager is one of the most important factors - if not the most important - of obtaining the performance, increase efficiency and effectiveness of the school organization[5].

Thus, to run a school organization is essentially to achieve, provided, decide, organize, command, coordinate and control in such manner that the system's features as a whole to achieve in the most favorable way: the school's mission, goals and objectives.

The pre- higher educational system ideal situation consists of the free, full and harmonious development of human individuality and to adopt a necessary scale of values for building an economy and a society of knowledge[6].

We must be aware that at pre-higher educational level the competitiveness is determined by the extent to which school organizations are able to develop these skills at young people who will then attend various components of training offer and academic training. It is essential for the school to determine scientifically whether the provided services arise to this mission and what has to achieve at all levels in order to remain or become competitive.

From the perspective of valorize the expert's competence to act, in this new context of knowledge-based society the investment in education is evaluated not only in curricular parameters but also through the outcome of the training process visible at the level of the skills acquired. All these are carefully monitored through the systemic perspective at all levels of the school organization.

The role of school's manager is essential, being the one that can convince the team members about the need of continuous improvement, the need of continuous improvement of the methods and

[4] Joiţa E., *Management educaţional,* Iaşi: Polirom, 2000;
[5] Ion Verboncu, *Ştim să Conducem?,* Editura Economică, 2005;
[6] *** Legea Educaţiei Naţionale nr1/2012;

resources used[7]. For management activity within the school may be proposed a range of skills, tasks and activities that meet specific educational activity in schools. The manager of the school should have the skills presented in Figure no. 3.

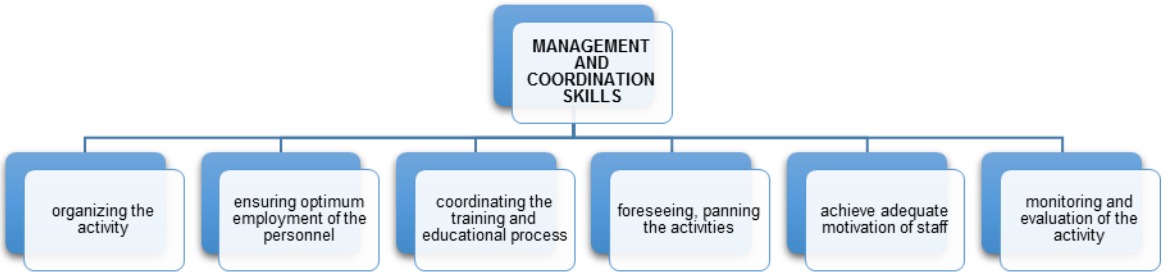


Figure no. 3 The management and the coordination skills

Through good communication, "the boss" and "the subordinate" agree on office issues and each must be receptive to the other's point of view. Thus, a competent and operative manager is not the one who does everything alone, but the person who manages to discover the skills of his subordinates, who knows how to delegate and supervise only the actions, the one who knows to build good work teams, using the intellectual capital and personal assets of the team members through decision making activities.

For educational organization which functions and acts in any social and economic system, the **decision** is the most essential act being the main tool for building the future, because is it important to decide what should be done today in order to have a future[8].

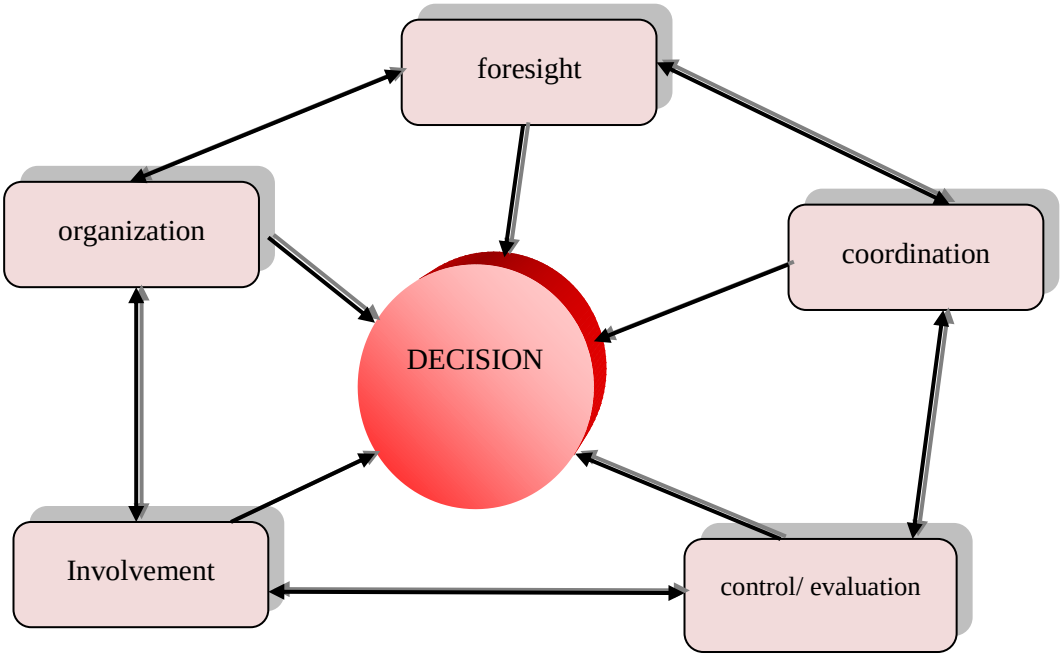


Figure no. 4 The relationship managerial decision-management functions

In any management process, and also in the school organization, the decision is the engine of any economic and social activity and not being a general function of management, is directly related to the performance of each of the management's functions and all elements that define the leadership (fig.no.4).

The decision, primary component of the ***decisional system***, is an essential element of management, stated by many authors, to be the most important specific instrument of expression. In fact, the quality of school's organization management it is manifested through the best designed and

[7] Dragomir, M., Pleşa, A., Breaz, M., Chicinaş, L., *Manual de management educaţional pentru directorii unităţilor de învăţământ*, a Centrul Regional de Dezvoltare şi Inovare a Resurselor din învăţământ Cluj, Ed. Hiperborea, Turda, 2000;

[8] Peter F. Druker, *Management strategic*, Editura Teora, Bucuresti, 2001,

implemented decisions[9]. The **decision** is the management's piece of resistance, its most active and dynamic expression through which its exercising its functions. Based on the data provided by decision practice, the analysis of different specialists' points of view has been formulated the definition of decision as being *the course of action chosen to achieve one or more objectives*[10].

The concept of managerial decision in school organization includes, cumulative, some defining ***conditions***[11]:

- *The existence of some extensive and reliable information,* in order to diminish the risk and the uncertainty of the decision maker. Decision can not be a random act, is a process of knowledge, a reasonable and volunteer to turn information into action through thorough study of resources, specific conditions and factors that may condition the conversion and expected results (fig. no. 5);

Figure no. 5: Factors that may condition the decision in school organization

- *The cluster of decision's alternatives from which it is operated the selection to consist of at least two elements*. Choosing from a number of solutions is a difficult operation; if there is only one possibility of action, the decision it is imposed by itself without the need of a choice;
- *The existence of a defined endpoint* (one or more targets achieved);
- *Application of decisional alternative chosen to influence the actions and / or behavior of at least one person other than the decision maker*.

If it wouldn't be respected one of the conditions listed could not take place a decisional process. The lack of accurate information may induce erroneous decisions. If there were no multiple variants for decision it could not be chosen the optimum variant. If there were no finality, it would not trigger the action to achieve the desired objectives. And in order that the decision to be of management it is necessary that its application to influence the activity and / or behavior of at least one person other than the decision maker. This is because the management process itself is an intellectual activity through which the manager determines that other people (subordinates) to develop activities that aim to achieve predetermined goals.

The decisional relations' characteristics at the school organization level are influenced and determined by the social -economic, technical -material and the human resource's level involved in the decisional process. The analysis of the decisional relations can be accomplished having in view the main variables that affect them (fig.nr.6). These relationships are conditioned by: the type of organization, its size, the characteristics of the educational process, material supplies, the degree of automation of information, human resources, the managers' view on school's management, educational law, etc. [12]

[9] Mihăescu Diana, *Modelarea deciziilor manageriale din instituţiile educaţionale-o abordare cibernetică*, Ed. Universităţii "Lucian Blaga" Sibiu, 2009;
[10] Nicolescu O., Verboncu I., *Management*, Ed. Economică, Bucureşti, 1995;
[11] Chivu Roxana, *Elemente generale de managementul educaţiei*, Ed Meronia, Bucureşti, 2008;
[12] Mihăescu Diana, *Modelarea deciziilor manageriale din instituţiile educaţionale-o abordare cibernetică*, Ed. Universităţii "Lucian Blaga" Sibiu, 2009

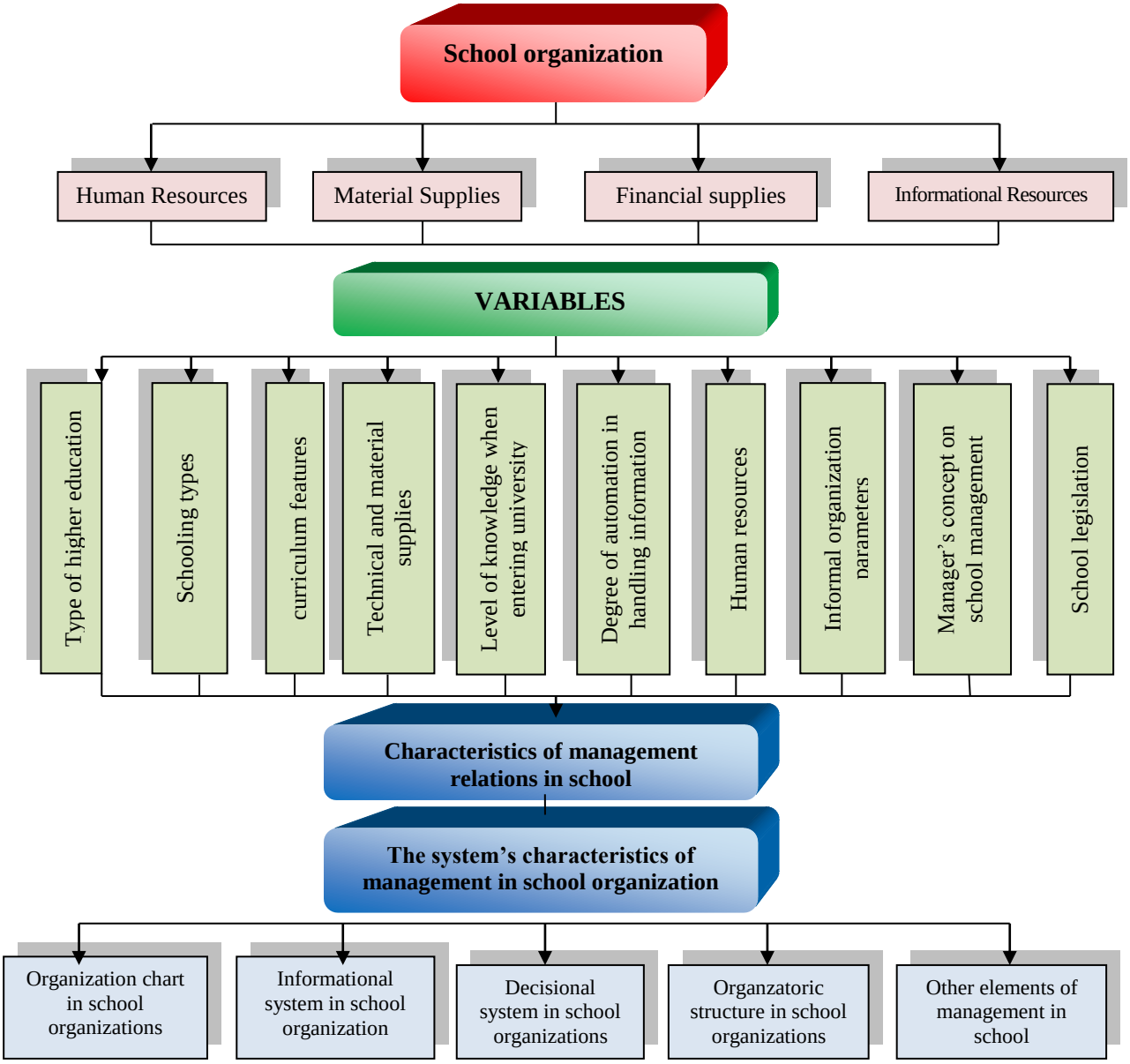


Figure no.6: Variables influence in decisional process relations at school organization level

The identified relationships at the school's organization levels are of several types, depending on the employees in a "partnership" its goal being to achieve the institutional objectives – pre-established goals.

In the school organization for a number of organizational purposes, structures and interactions (relations actually inside or outside the formal structure). The function of the school organizations in the system's framework is to make the "entries" (people - students, teachers - financial, informational, educational, cultural, social) in "outputs" - the results of education and training activities: "educated" people, capable of social and professional insertion and social engagement, with a high level of culture able to "create value".

A crucial resource for developing the decision-making process and for the quality and performance of the process management at the level of school's organization is the information, and the information system (fig.nr.7). The basic elements of an *informational decisional system* in schoolorganization are: the data and the information; material aids of information, transmission

channels or informational circuits, processing procedures, means of processing informational flows[13].

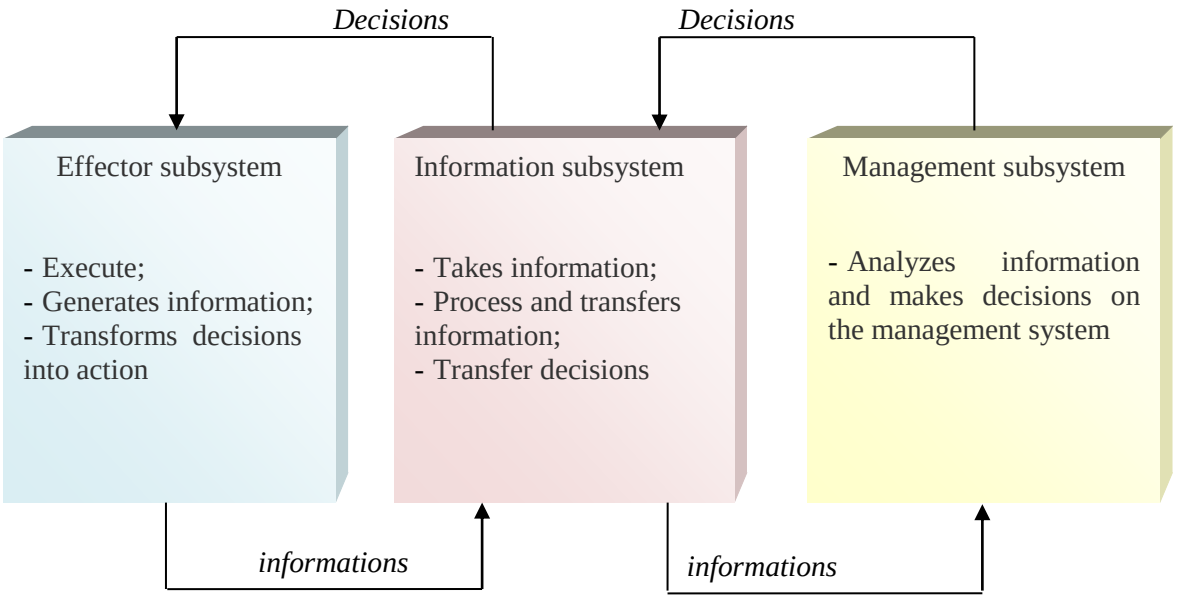


Figure no.7: Structure's pattern of the informational process and decisional at the school organization

The mission of the school organization it is the main task of the employees in the educational process and is based on good management of resources of all kinds - curricular, informational, human, financial, etc.

The relations at the organizational level and optimize their procedure is for the school manager a priority because it is essentially depending the mission's fulfillment. Given the managerial functions, the pre-university manager assigned the tasks not only for planning, implementation, monitoring and evaluation, but also tracking the effectiveness of the functioning the relations at the organizational level.

The practice in school organizations shows that school manager in professional work a school organizations, are increasingly using the management's theory and methodology solves the situation, based on analysis of various data, communicates argumentatively the decisions entrusts responsibilities, negotiate conflicts, accept errors as possible, and also the initiatives, evaluates and appreciates each member of the school's performance after the results, relates and exploits every opportunity that leads to the affirmation of the school's organization[14].

Therefore, the school organization manager is an important decision factor who through science (knowledge management, psychology, sociology, law, accounting, finance, etc..), through personal qualities and managerial skills manages effectively and ensure stability of the school organization and the necessary progress.

[13] Liviu Mihăescu, *Sisteme informaţionale şi aplicaţii informatice în administrarea afacerilor*, Editura Universităţii „Lucian Blaga" din Sibiu, 2009;
[14] Cojocariu Venera, *Introducere în managementul educaţiei*, EDP, Bucureşti, 2004;

TABLE OF CONTENTS