Volume 31 Number 4 December 2007 ISSN 0350-5596

Informatica

An International Journal of Computing and Informatics

Special Issue:

eGovernment in Balkan Countries

Guest Editors:

Matjaž Gams
Marjan Gusev
Božidar Kliček
Veljko Milutinović
Brane Vujičić

EDITORIAL BOARDS, PUBLISHING COUNCIL

Informatica is a journal primarily covering the European computer science and informatics community; scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers accepted by at least two referees outside the author's country. In addition, it contains information about conferences, opinions, critical examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and information industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor from the Editorial Board can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the list of referees. Each paper bears the name of the editor who appointed the referees. Each editor can propose new members for the Editorial Board or referees. Editors and referees inactive for a longer period can be automatically replaced. Changes in the Editorial Board are confirmed by the Executive Editors.

The coordination necessary is made through the Executive Editors who examine the reviews, sort the accepted articles and maintain appropriate international distribution. The Executive Board is appointed by the Society Informatika. Informatica is partially supported by the Slovenian Ministry of Higher Education, Science and Technology.

Each author is guaranteed to receive the reviews of his article. When accepted, publication in Informatica is guaranteed in less than one year after the Executive Editors receive the corrected version of the article.

Editor's Introduction to the Special Issue

eGovernment in Balkan Countries

Western Balkan Countries (WBC) were involved in the most conflict events in recent European history. Former Yugoslavia, consisting of Slovenia, Croatia, Bosnia and Herzegovina, Serbia, Montenegro and Macedonia ceased to exist. At this point, Slovenia has already joined the European Union including Euro and the Schengen borders - border crossing without any border control post, just the common EU-state sign welcoming the visitor. It is believed that WBC will soon or later follow.

There is a strong determination in the European Union and in Western Balkan Countries that WBC join EU. To foster the integration proccess, the European Union started several projects, including coordination actions (CAs) as part of the IST call 6. Among the accepted CAs, one represents the central part of this special issue: Facilitating eGovernment Innovation and Enabling Project Implementation in Western Balkan Countries (We-Go).

The idea is to introduce EU standards in regard to eGovernment in the past years. Obviously, action plans for eGovernment in local information system strategies are not fully realized and the systems are not conformant with EU policies and requirements. The low level of awareness of eGovernment services is an additional hurdle towards EU conformance. In order to ensure sustainability of eGovernment in WBC the WeGo proposal seeks the following major actions to be taken:

- Firstly the introduction of an eGovernment Framework and Interoperability Platform with focus on transactional (cross border) services, and providing context-rich descriptions for selected application domains as best practices, including demo / trials assessments for facilitating and enabling follow-on implementation projects including relevant specifications.
- Secondly, the establishment of specialized eGov Academies will increase the level of awareness regarding eGovernment and enhance the management capabilities for eGovernment solutions. Using the knowledge and input of EU government approaches the academies will deliver contemporary and actual knowledge by offering trainings and certifications for the different target groups.
- Thirdly, the eGov Resources Network Approach and the eGov Innovation Progress and Impact Evaluation will ensure sustainability and further improvements.

The project builds up from recommendations on how to improve eGovernment in the WBC countries by using scientific and practical best practices from EU and action plans by WBC governments.

If the actions mentioned are to be realized, the WBC countries and their eGovernment efforts could move a major step closer towards the EU set in i2010. The realization of action plans will support WBC countries in establishing and integrating eGovernment solutions. With measures and actions created, the EU could help the WBC countries to evaluate the efficiency of their activities and guarantee sustainability.

The current action plans in the WBC are targeted to meet the requirements set out in eEurope 2002 and eEurope 2005 and most of them are not met yet. Meanwhile we are facing the additional challenge to meet criteria set out in i2010. Therefore tremendous effort is needed to accelerate the execution of relevant action plans and to achieve EU alignment. Current overview indicates that:

- Most of the countries have adopted fragments but not set up the framework needed for eGovernment legislation. We conclude that in the WBC countries a comprehensive legal framework for eGovernment is missing.
- There is lack of interdisciplinary approach and cross-organizational initiatives as well as management capabilities to allow the management of such eGovernment projects.
- The level of IT knowledge and IT skills in regard to eGovernment for developers and information officers needs to be increased.
- Concerning applications for eGovernment basic building block e.g. registries are not fully implemented or completely missing. For cross-border and legal certainty purposes within the common EU market the implementation of such eGovernment applications has a high priority.
- The ICT infrastructure has isolated "island" structure, realized on heterogeneous platforms without possibility for interoperable functioning - individualidual solutions, no interchange of data and information, communication and transactions. They do not address regional aspects or interoperability compliance with the EU standards.
- The digital divide gap, content gap, gender gap, multicultural and multilingual gap in WBC is increasing in comparison to the EU due to following facts:
  - lack of available e-services and
  - lack of ICT knowledge and skills
  - lack of possibilities for democracy and citizen inclusion including multicultural and multilingual aspects.

The project We-Go is aiming to transfer and to adapt successfully "Good eGov Knowledge and Practices" and to create a sustainable innovation impact within the WBC Countries by facilitating and enabling imple-

mentation projects and aligning to EU standards set in eEurope 2005 and i2010.

Two of the partners in the WeGo project have achieved significant progress in recent years. Currently, Austria is the first European country in recent polls regarding eGovernment activities and Slovenia is second to third. It is expected that these countries together with other EU partners will be able to help improve progress in WBC in particular due to traditional cooperation.

In this special issue, overviews of current state of the art in Slovenia, Croatia, Bosnia and Herzegovina, Serbia and Macedonia are presented with the aim to further promote and advance eGovernment in Western Balkan Countries.

Matjaž Gams, Slovenia
Marjan Gusev, Macedonia
Božidar Kliček, Croatia
Veljko Milutinović, Serbia
Brane Vujičić, Bosnia and Herzegovina

eGovernment in Slovenia

Aleš Dobnikar and Alenka Žužek Nemec
Ministry of Public Administration
Republic of Slovenia

Keywords: eGovernment, Slovenia

Received: June 12, 2007

In Slovenia, like elsewhere in the European Union, the eGovernment is one of the features of the modern face of the public administration. The other features at the forefront of the makeover are the reduction of administrative barriers, the improvement of services offered by the public administration (and subsequently the satisfaction of its users), the introduction of quality management and business excellence tools, and human resource management. What we are doing in Slovenia is by no means a reform of the public administration, as there are no revolutionary changes. Instead, we prefer to speak of a process of perpetual improvement, as that applied in successful businesses, with the aim of constantly improving operations.

The eGovernment project in Slovenia represents the use of information and communication technologies as a tool to making the public administration better. Technology can help the public administration function more transparent and make it easier to manage, reducing its operating costs and improving the quality and accessibility of its services.

Povzetek: Podan je pregled e-uprave v Sloveniji.

1 Introduction: eGovernment as a tool for public administration modernization

The modernization of public administration is not the primarily issue and political priority just in Slovenian government, who for this task established the new Ministry of Public Administration [1], but is very important issue also in all other national and common EU levels. The modern technologies, with all today's available electronic means, is very important mechanism to provide modern public administration, however definitely not sufficient. We strongly believe that only conjunction of simplification of processes, reducing administrative burdens and in-advance impact assessment, strongly connected and supported with information and communication technologies give the appropriate synergy to develop modern public administration.

The modern and democratic public administrations have all the time maintain the trust of its citizens through the following visions:

- orientation of public administration towards users, what puts the user in the centre of the process and the provision of an efficient and competitive servicing of individuals, civil society and, in particular, of the Slovenian economy.
- a high-quality and efficient functioning of the public administration, which includes the establishment of a quality system in the public administration, permanent tendency to increase efficiency and quality at all levels of the governmental and administrative decision-making with minimum backlogs, and based on a pragmatic and a rational functioning.
- openness and transparency which covers an easy, comprehensive and free access to information of public nature and the provision of the co-operation of the general public in decision-making.

1.1 eGovernment for better public administration

eGovernment tries to fulfil all these elements that define modern public administration. The use of information and communication technologies should provide kind, simple, accessible and safe electronic administrative services and information which shall be all the time available via internet to individuals and companies in all life situations as well as for internal government administration.

Slovenia is putting special emphasis on ensuring that state institutions do not demand from clients information that is available in public registers. This principle is written down in the Act on administration procedures and Slovenians are increasingly aware of their rights in this respect. Of course, this principle makes life harder for the state institutions, extending procedures when the information is not available online. Well-maintained, up-to-date and fully integrated databases that can be easily accessed by all state institutions are essential for the good functioning of the public administration in offering fast, simple and efficient services. A suitable system of interoperability ensures that the data is collected only

once and is made available to those who need it. It also enables the integration of data services for online public services for individuals and companies.

Today's technologies enable us to put into practice in public administration virtually all ideas, documents or forms, we can have free access to the governmental services, we could participate in decision-making procedures, technology can fulfil highly demanded security for privacy and for identification. Definitely, today's information and communication technologies and eGovernment services are indispensable elements on the path for achieving better public administration.

1.2 Importance of ICT in EU

Information and communication technologies are a powerful driver of growth and employment. A quarter of EU GDP growth and 40% of productivity growth are due to ICT. Differences in economic performances between industrialised countries are largely explained by the level of ICT investment, research, and use, and by the competitiveness of information society and media industries.

Electronic invoicing in Denmark for example saves taxpayers €150 million and businesses €50 million. If introduced all over the EU, annual savings can amount to over €50 billion.

In the next sections the eGovernment in Slovenia will be presented in more details. First, the main strategic documents, coordination and regulatory framework will be provided. Next, the overview of the main developments will be presented.

2 eGovernment profile in Slovenia

In Slovenia the Ministry of Public Administration is responsible also for eGovernment [1]. The ministry was set up in 2004, incorporating all offices whose common goal is to improve the public administration. This step means efforts to improve and simplify public administration procedures and to develop eGovernment under one roof and one leadership. The ministry firmly believes in the success of eGovernment and is confident that online will become the preferred way of doing business with state institutions in the coming years.

In the past years several eGovernment applications have been developed and integrated within the own area of competences, for example taxes. Nevertheless, steps are being taken towards horizontal integration covering several departments and institutions.

The current strategic framework for the development of eGovernment in Slovenia is comprised of four key documents:

- Slovenia's Development Strategy, adopted by the Government in June 2005;
- The eGovernment Strategy of the Republic of Slovenia for the Period 2006 to 2010 (SEP- 2010 "eGovernment for effective public administration"), adopted by the Government in April 2006;
- The Action Plan for eGovernment for the period 2006 to 2010, adopted in February 2007;
- Strategy for the development of the Information Society in the Republic of Slovenia until year 2010 (SI2010), adopted in February 2007.

In 2006, Slovene eGovernment entered a new period of development with effective and visible new results from the period of the first eGovernment strategy in public administration and other influential strategies and programmes up to 2006. According to various criteria it has achieved a level of development comparable to, or higher than that in other EU states. It has established an efficient and reliable information and telecommunications infrastructure for electronic services for citizens, other natural persons, businesses, other legal persons and public employees. The eGovernment portal (»eUprava«) and others offer information and electronic services; the results of their use are visible, which encourages further work in this area. The numerous measurements, comparisons and results are at this moment entirely encouraging, but are not the most important factor. The most important factor or global objective, and one which will be the primary driving force behind the further development of eGovernment, are satisfied users. Their satisfaction can be achieved through friendly, accessible, simple to use and affordable electronic services

In April 2006 the Slovenian government adopted an eGovernment strategy for the period from 2006 to 2010 [2]. The purpose of this new strategy is to lay down a framework and goals for further realization of new and existing eGovernment activities, especially regarding user satisfaction, rationalization of administration, and modern e-services, which will raise the quality of life and make contact with the administration more user-friendly. The strategy takes into account modern guidelines and initiatives which have been passed at EU level and which lead to success throughout the EU. It also takes into account the initiative "i2010 – A European Information Society for Growth and Employment" and the Ministerial Declaration and guidelines from the ministerial conference "Transforming Public Services, 24 November 2005, Manchester, UK".

The strategy sets out the following eGovernment targets to 2010:

- focus the operations of public administration on user needs;
- increase the quality and efficiency of the functioning of public administration;
- increase user satisfaction;
- reduce administrative burdens;
- increase the transparency of operations of public administration;
- achieve synergetic effects at all levels of public administration through the use of eGovernment;
- include the widest circle of users in the decision-making process (e-Democracy);
- reduce the burden on human resources in administrative procedures.

Numerous projects and activities have been started and will be carried out in order to implement the above

targets through different periods. The most significant actions foreseen include, per period:

- In 2007 and 2008 major progress will be made towards more efficient internal administration operations with support of eGovernment. The emphasis will be on upgrading and integration of internal administration processes, introducing standardised horizontal and vertical solutions, implementation of standardised information technology platforms, development of interoperable eGovernment solutions and services, new operational models, user training and establishment of uniform eGovernment architecture. The revision of business processes and linking of administration databases are especially important to the success of eGovernment. Business processes must be efficient, transparent, interconnected and oriented towards the needs of users of eGovernment. Revised business processes form the basis for achieving the objectives of eGovernment.
- In 2009 and 2010 all three branches will be better connected. Efforts in this period will be directed towards complete integration of the information and telecommunications infrastructure, revision of business processes at the level of the entire eGovernment and the integration of various e-services into a unified system for complete processing of life events of citizens and businesses.

A central element of the eGovernment strategy is the Action Plan for eGovernment for the period 2006 to 2010, adopted in February 2007 [3]. The primary aim of the Action Plan is to give concrete form to the implementation and monitoring of the eGovernment strategy in Slovenia. It describes in a more detailed manner the actions announced in the strategy for reaching the targets set for 2010. The Action Plan also provides detailed updates on the progress made so far. It also includes examples of good practice and a general overview of the advancement of eGovernment in Slovenia, comparing it with progress in other parts of the European Union. To ensure transparency, particular care was taken in selecting the most appropriate processes, organisational solutions, tools, IT solutions and methodologies to be used throughout the plan. These are essential for the successful development of eGovernment in the future.

A secondary goal of the Action Plan is to give new impetus to the development of e-Services that are considered necessary but that have so far been delayed. In particular, those based on joint EU projects, using a shared architecture and common European standards, as well as those necessary for the internal functioning of the Slovenian Government.

With regard to the eGovernment strategy's objective of user-oriented public administration, the Slovenian Government has started implementing a set of reforms under the slogan "friendly and effective public administration". Among planned measures and objectives are the creation of a business-friendly public administration by consolidating services in line with the one-stop-shop principle, the elimination of unnecessary red tape to public procurement and the streamlining of the public sector. Furthermore, the government has adopted a 'Programme of Measures for the Reduction of Administrative Burdens' (November 2005) which contains over thirty concrete measures aimed at simplifying procedures and raising the quality of public services (see section »eGovernment and reduction of administrative burdens" for more details).

2.1 eGovernment regulatory framework

In Slovenia currently there is no overall eGovernment legislation. Special arrangements can be still provided by jurisdictional field regulation; however, the mainstream legislative approach is that only few differences to the general provisions of the General Administrative Procedure Act are regulated by jurisdiction field specific legislation, none of those derogating from the concept of electronic commerce in administrative proceedings.

However, an Act on Interconnection of Public Registers is in preparation, which will set a framework for the development and implementation of e-Public services.

Moreover, in 2006, the Slovene government adopted the General Administrative Procedure Act (Official Gazette of the Republic of Slovenia, no. 105/2006-ZUS-1) which provides the general legal basis for all administrative proceedings; i.e. all A2C and A2B and a major part of A2A relations. Among the main provisions of the new Act is one allowing for two-ways and full electronic communications between administration and citizens. Before the entry into force of this text, citizens could post their e-Documents though the e-Services on the eGovernment state portal by using the web application and digital signature, but the answer from the administration could be expressed by classical mail only. This Act thus legalised what is qualified as "e-Serving".

Other important legislations have to be considered for the use of electronic documents and electronic services:

- Personal Data Protection Act (Official Gazette of the Republic of Slovenia, No. 86/2004, 113/2005-ZInfP). The main goal of the Act is to prevent any illegal and unwarranted violations of personal privacy in the course of data-processing, and to ensure the security of personal databases and of their use.
- Electronic Commerce and Electronic Signature Act (Official Gazette of the Republic of Slovenia, No. 57/2000), as a horizontal bill regulating e-commerce in a broad sense applies also to administrative, judicial and other similar procedures unless otherwise provided by another law. It provides the legal basis for using e-signatures.
- Decree on Conditions for Electronic Commerce and Electronic Signing (Official Gazette of the Republic of Slovenia, No. 77/2000 and 2/2001) defining in detail individual conditions from the act, prescribing special, rigorous conditions regarding Certification

Authorities, who issue qualified certificates (compulsory liability insurance, special requirements regarding equipment and employees, exacting procedures, internal regulations, etc.).

- Electronic Communications Act (Official Gazette of the Republic of Slovenia, No. 43/2004, 86/2004-ZVOP-1, 129/2006, 102/2007) Its aim is to establish effective competition in the electronic communications market, to maintain effective use of the radio frequency spectrum and of the number space, to ensure universal services and to protect user's rights.
- Protection of Documents and Archives and Archival Institutions Act (Official Gazette of the RS, No. 30/2006). It regulates electronic content management. All electronic records (including electronic documents) shall have full legal effects under some technical consideration. The act also regulates registration and accreditation of equipment and service providers.

3 Main developments and key milestones

Over the period 2001 to 2006 the Slovene public administration underwent several organisational changes. EGovernment developed in parallel and in accordance with them and with the existing possibilities, and was marked in that period by strategy and programme documents such as eGovernment strategy until 2004, eGovernment Strategy for Local Self-Government, the Strategy of the Republic of Slovenia in the Information Society and others. Numerous successes were achieved up to 2006, while at the same time all participants faced new challenges and failures within eGovernment projects. A critical assessment, the sum of all of our experiences and acquired knowledge, represents the driving force for the new era of the development until 2010. In the following the main developments are presented.

3.1 State Portal for entrepreneurs (One stop shop) or e-VEM

The e-VEM project follows the strategic goals of the Ministry for Public Administration and the Government of the Republic of Slovenia. The project follows those public administration's goals that are focused on users and creation of a friendly environment for the development of entrepreneurship. The project has started on 1st July 2005 and the product of the project is a State Portal for Legal persons, e-VEM, http://evem.gov.si (see Figure 1). The project is covering two scopes of activities and relationships: government – legal persons (G2B) and government – government (G2G). The project is based on a one stop shop concept, because procedures can be electronically done at one spot, we also call it e-one stop shop.

The basic purpose of the e-VEM project is to provide

Figure 1: State Portal for entrepreneurs.

a suitable information support for the future entrepreneur and enable him/her to start with business operations in the shortest time possible. The information support provides a unified support regardless of the type of entrance into the system. The support is the same for the submission of electronic application for registration of a future entrepreneur via internet as well as for the submission of application, which has been made for the entrepreneur by an advisor (person) that is offering support and help to the entrepreneur on one of the local entry points. The e-VEM Project offers information support to all enumerated entry points. This way the unification of the procedures is achieved and all information is gathered at one spot.

All of these services can be accessed from home by using digital certificates provided by one of the certification authorities in the Republic of Slovenia, or in person at any of the more than 200 VEM access points in the country. The aim is to enable future entrepreneurs to complete all formalities required for establishing a company in one place [6]. The Ministry of Public Administration is planning to expand the e-VEM system to facilitate registration of other types of companies (ltd.,...) from the beginning of the 2008.

The e-VEM portal brings considerable savings for entrepreneurs registering for the national register of independent entrepreneurs due to its "one-stop-shop" nature and the cancellation of registration and other fees. The statistics shows that the number of new registered natural person has increased for 21,7 % since the project started in July 2005. We have abolished paying tax for registration; a natural person pays nothing to open a business, to make changes in the register of companies, to close the business. We have reached higher interest among citizens to start running a business. The numbers show 21,7% higher number of registration from the time of the production of the system and 54,4% more sole traders operating in comparison with the year before the introduction of the system e-VEM. We have estimated savings of citizens. Sources of savings are: elimination of paying fee for registration, closure and changes, elimination of fee for forms, reduced costs of transport and time due to one stop shop concept. Estimated savings at the beginning of the project amounted 766.667,00 EURO per a year.

3.2 e-Tax System

Slovenian e-Tax system is a complete business solution combining a web portal with back office integration and the highest level of security based on a PKI infrastructure. It connects also other governmental institutions and includes information exchange as required by EU regulations, in particular the E-Commerce Directive [7].

The case provides an example of how a transactional eGovernment service can provide fast, accurate and secure tax reporting for citizens and companies. It also shows how new efficiencies were created for the Tax Administration by speeding up the processing of tax returns, cutting down on paper, reducing errors due to data re-entry and improving employee productivity. The system allows individuals and companies to file taxes online using a qualified certificate issued by any registered certification authority in the country. In this way, the solution helps to increase the take-up and use of digital certificates for not only tax filing but for other public and private purposes as well.

3.3 Land register

The land register is a public register kept by courts and containing data on real property rights. Its basic function is to make public the information on rights and legal facts relevant to legal relations regarding real estate. To access it, users must first register.

3.4 e-Cadastre

Through the web, users can access data from the land and property register, kept by the Surveying and Mapping Authority of the Ministry of the Environment and Spatial Planning. This register dates back to the days of Empress Maria Theresa of Austria-Hungary.

3.5 Data exchange within public administration

Some of the central registers:

- Permanent population register,
- Central population,
- Register of birth, marriages and deaths,
- Tax register,
- Business register,
- Court register.

Allow for quick and efficient exchange of data between state bodies. Citizens are thus relieved from having to provide certificates, copies and other documents.

3.6 Online car registration

This service enables to extend the validity of your vehicle registration certificate, which simultaneously extends the registration of your vehicle. This service takes advantage of e-business and uses e-applications, e-payments and e-signatures. To extend the validity of a vehicle registration certificate through the web, you only need the vehicle registration certificate number, the valid mandatory insurance policy number and a method of payment. You will receive the new certificate to your mailing address.

To register a car, individuals must insure it and pay road tax. Both can now be done online, while the individual receives the documents (unfortunately hard copies of documents are still in use in Europe) through the post.

3.7 Notification of official documents expiry

The eGovernment portal gives individuals the option of being notified when their documents expire. Using a digital certificate, they can access a list of all their

documents (passport, ID card, driving license, gun permit, etc.) and select for which they want to receive notification through the post of impending expiry. Individuals can also change their place of residence and name and access also other services online.

3.8 Court register

This is a public register containing information on companies (amount of capital, representatives, etc.), which is relevant to legal relations. A digital certificate is not necessary to access this register.

3.9 ISPO public administration data

The ISPO system allows viewing of public economic data (imports, exports, price indexes) and administrative data (workforce, wages) and displays them dynamically.

3.10 View your own personal data

This feature provides citizens with secure and traceable access to their own personal data in the Central population register. A governmental digital certificate is mandatory.

3.11 e-Democracy

The essence of e-Democracy is in the use of new technologies to enhance and encourage democracy and politics as part of people's lifestyle. Citizens are thus involved in the process of forming government politics, decision-making processes and legislative procedures (Figure 2: the example of published legislation on the National Assembly web-site).

3.12 Public information catalogue

Represents a collection of public information data, classified by topic, and available to the competent authorities. The user can access public information all in one place – the procedure is fast, transparent and simple.

4 Main eGovernment infrastructure components

4.1 The State Portal of the Republic of Slovenia

The eGovernment portal "e-uprava" was launched in March 2001 and re-launched in December 2003 and modernised in May 2006. The enhanced portal supports

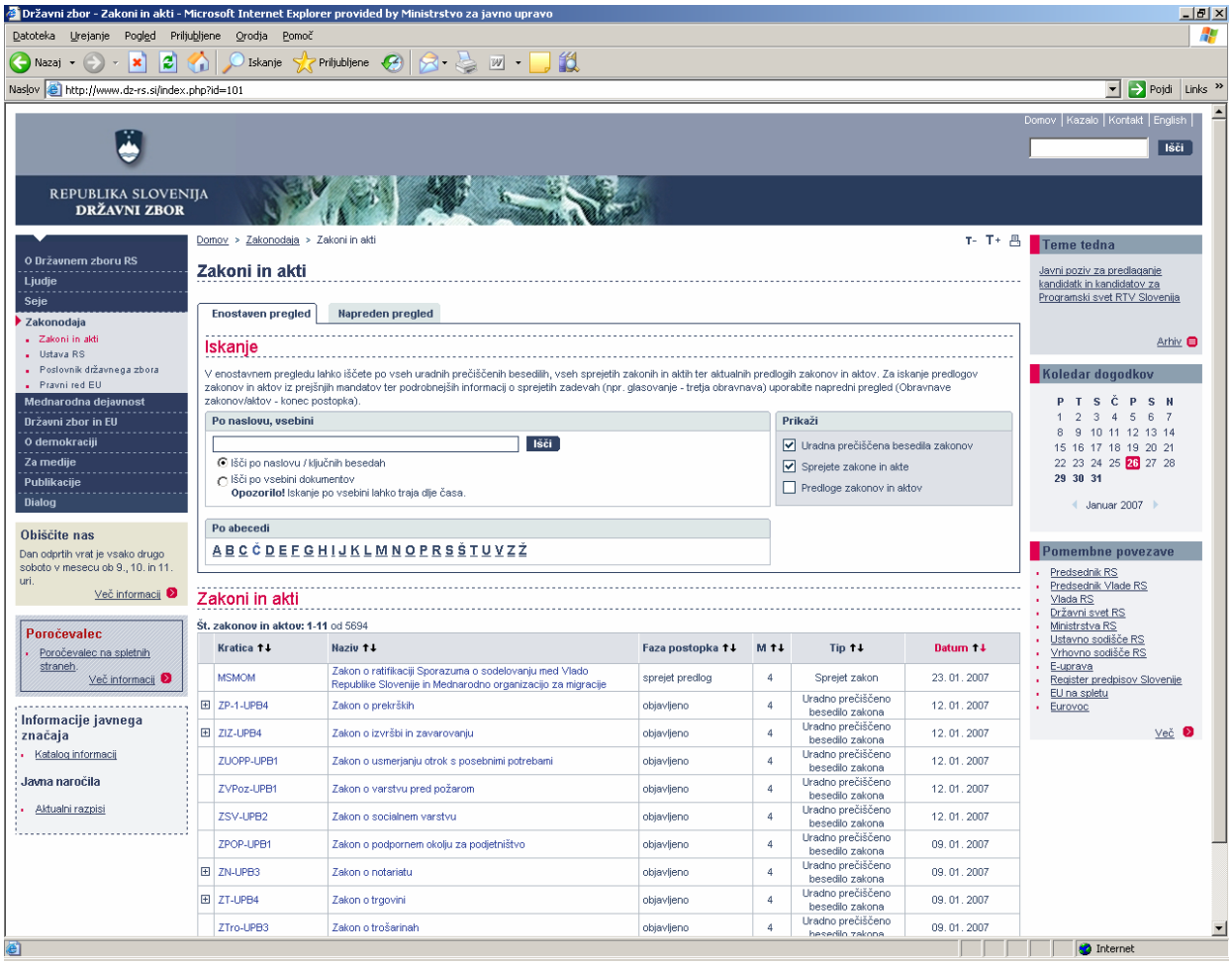


Figure 2: All the legislation (including the proposals) is published on the website of the National Assembly (http://www.dz-rs.si/).

Government to Citizen (G2C), Government to Business (G2B) and Government to Government (G2G) interactions and offers various services to citizens, legal persons and public employees. The portal provides access to the Electronic Administrative Affairs application, which supports full electronic handling of administrative forms registered in a centrally maintained registry of procedures. The application could be used by all residents equipped with qualified digital certificates valid in Slovenia.

4.2 e-SJU portal

"e-SJU" portal stands for "electronic services of public administration" portal. The e-SJU portal represents a new part of the renewed state portal eGovernment. The portal e-SJU offers a single access point for all forms that can be published on the web by any public administration institution. The forms are published in different formats, which can be filled in by citizens and sent by them to the selected institution. The system includes the description of over 400 different services and 350 forms, but not all public administration institutions are included in the system yet. The goal is to have all public administration institutions in Slovenia participating in the system in the future.

4.3 Network

HKOM (Fast Communications Network) represents internal governmental communication network. Most government bodies have internet/intranet facilities and are linked to a government-wide network HKOM, connecting more than 1.600 local computer networks. E-Identification and e-Authentication infrastructure (Public Key Infrastructure - PKI) has been deployed in Slovenia and four certification authorities (Certificate Services Providers – CSPs) have been accredited: the Ministry of Public Administration (SIGOV-CA for government communications and SIGEN-CA for the general public), HALCOM-CA, AC NLB, and POŠTA CA. Slovenia adopted the EU Directive on electronic signatures by the Act on Electronic Commerce and Electronic Signatures (ZEPEP) in the year 2000. In 2004, a further act amending the Act on Electronic Commerce and Electronic Signature entered into force in order to create a legal basis for an upcoming e-Identity Card project.

5 eGovernment and reduction of administrative burdens

The eGovernment project in Slovenia represents the use of information and communication technologies as a tool to making the public administration better.

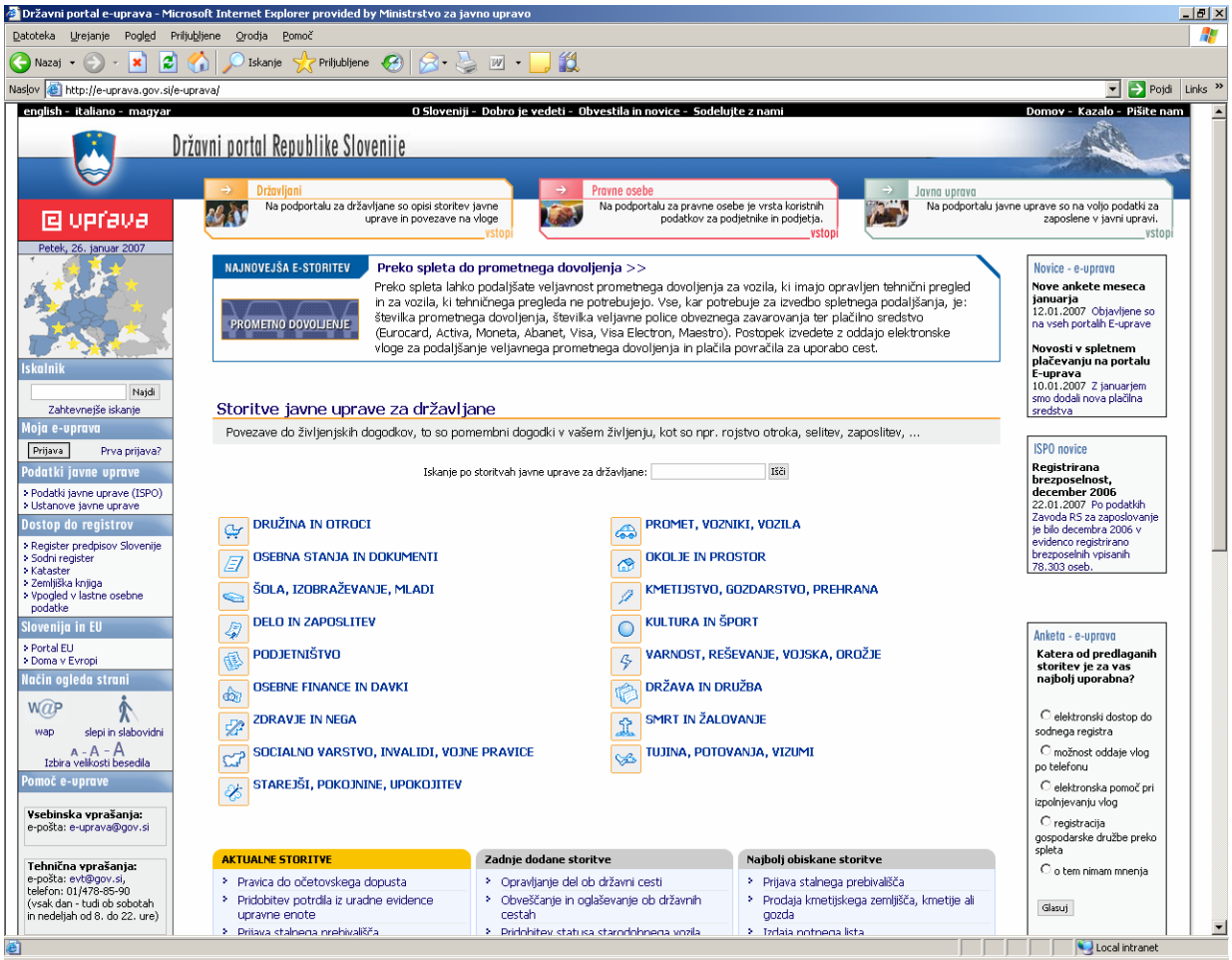


Figure 3: The State Portal of the Republic of Slovenia: eGovernment for all in one point – citizens, businesses and public sector

Technology can help the public administration function more transparently and make it easier to manage, reducing its operating costs and improving the quality and accessibility of its services.

The eGovernment project is closely linked to the overhaul of administrative procedures. Digitalisation of services and processes breeds rationalisation by prompting you to analyse the process and make it better by eliminating the excess, valueless elements. If nothing else, the process must be standardised prior to taking digital form so as to ensure uniformity at all levels. A service must be fully standardised once it is placed on a portal.

One must keep in mind the fact that the eGovernment is by no means purely a technological matter, but rather a makeover of the organisation and processes involved. For this reason the eGovernment is closely associated with a project on the mind of virtually every government in the European Union – that to reduce administrative barriers. In recent decades, a mountain of irrational and valueless procedures has piled up around Europe, wasting citizens' time and money. This bureaucratic load was a side effect of public policies, which failed to pay any attention to the issue. The people making the regulations were given a free reign - the outcome was felt most by businesses, which lost their competitive edge. It is therefore not surprising that one of the fundamental elements of the Lisbon Strategy involves the simplification of procedures and improvements of legislation. Slovenia launched a systematic effort to reduce administrative barriers in 2005. For this undertaking the Ministry of Public Administration put together a team of dedicated and experienced professionals. The government made a commitment to require of all proponents of regulations (ministers) efforts to ensure that new rules do not cause new administrative hurdles. A report on alleviating administrative barriers is now a mandatory part of every new law and implementing regulations. This report is not merely another form, but a commitment to a new way of thinking for the makers of new regulations. In drawing up new regulations, state institutions must pay attention to alleviating existing administrative barriers and ensuring that no new ones are created. A team at the Ministry of Public Administration has the task of reviewing all regulations put up for government debate and if need be, rid it of unnecessary bureaucracy. The team also ensures that legislation in the making is brought to the attention of public stakeholders in cases where the proponent has not done so. This process has helped prevent the creation of at least ten new administrative barriers and has allowed stakeholders to participate in the making of laws even if the proponent had “forgotten” to include them.

In addition to the preventive measures, efforts to remove administrative barriers include the constant collection of proposals for simplifications. The have been were sent by public stakeholders, including various organisations (such as business chambers), as well as companies, citizens, public servants and their managers. Theirs and other proposals form the basis for an annual action plan for reducing administrative barriers that the ministry puts to the government [8]. The action plan is adopted early in the autumn for the coming year and commits all ministries to tangible changes in regulations. Nearly 100 simplifications have so far been carried out on the basis of such an action plan. One of the most prominent solutions is the "one-stop shops" for small companies. Instead of being given the run-around, an entrepreneur can now sort everything needed to set up a business in one spot, also via the Internet, in as little as two hours, without the costs and the complicated paperwork. A similar approach was introduced in November of this year for large corporations.

The action plan has also led to the abolishment of many reports and forms, the lowering of the criteria for opening a business, simplification of spatial planning procedures and elimination of some permits. Another feature means that Slovenian citizens no longer have to fill out personal income tax forms - as of next year, the Tax Administration will carry out all the work for them.

6 Summary conclusion and future trends

A relatively small public administration is Slovenia's main competitive advantage. Its small size allows for single solutions and facilitates coordination. We are committed to the principle of one-stop shops: all the information and public administration services are assembled in a single eGovernment portal. The portal was set-up as an online shop window, where all state institutions can showcase their services. Any portal should be more of an organisational achievement than a technological one, with the aim being to organise the content in a suitable and user-friendly way. The essence of every portal, of course, is the services it offers to citizens and companies. The electronic services must pursue two general principles: the principle of high value added and the principle of user-friendliness.

It is important to note that the eGovernment is closely linked to the optimisation of processes. EGovernment is by no means just a "fancy dress" wrapped in an appealing portal offering services. A key factor in achieving its success is the functioning of back-office systems. Take the one-stop-shop, for example - here the great majority of entrepreneurs are not opting for online services (electronic transactions make up only about 6 per cent of total transactions), but sort out their business at the offices of one-stop-shops. In either form, the quick and easy service is offered to the entrepreneur with the help of an information system running in the background. All the information related to the client is collected in one place and electronically distributed to the relevant institutions (Business register, Tax Administration, Pension Insurance Institute, Health Insurance Institute and so on). The information held by the state institutions is stored in databases. Broad information allows for quick decisions. All the required procedures are usually completed in the same day, allowing the company to begin operating virtually immediately.

Slovenian suggestion for good practice development of eGovernment is based on the following principles:

- one-stop access to services
- life situations approach
- focus on high impact services
- simplicity and user-friendliness
- mutual recognition of all digital certificates
- back-office is at least as important as front-office
- raising the importance of e-Participation and e-Democracy

The latest results of the survey of the European Commission on the development of e-services in across 31 countries (the 27 EU Member States, plus Iceland, Norway, Switzerland and Turkey (EU27+)), which is done annually by an independent organization CapGemini, by measuring 20 services for citizens and 12 services for companies show Slovenia's progress from 7th place in 2006 to the excellent 2nd place in 2007, which we share with Malta [9]. This achievement is a big acknowledgement for Slovenia and its continuous efforts to achieve a more effective public administration.

Regarding the future developments the focus in Slovenia like other EU countries is oriented in eGovernment contributing to political priorities of the European Union [10]:

- the reduction plan of European Commission – 25% set burden reducing targets,
- eGovernment services supporting climate protection,
- solutions for Inclusive eGovernment/e-participation with taking into account same commercial standards to reach the same quality as online business services,
- multi-channel approach,
- taking on-board new trends and new technology.

References

[1] Ministry of Public Administration Republic of Slovenia, http://mju.gov.si/

[2] eGovernment strategy for the period from 2006 to 2010, http://www.mju.gov.si/en/legislation/-important_documents/

[3] Action plan for eGovernment in Slovenia, http://e-uprava.gov.si/eud/e-uprava/akcijski_nacrt_e-uprave_2010_1.pdf

[4] eGovernment Strategy for Local Self-Government, (in Slovene only) http://www.svlr.gov.si/si/-delovna_podrocja/podrocje_lokalne_samouprave/informatizacija_obcin/

[5] Information Society Strategy (in Slovene only) http://www.mvzt.gov.si/fileadmin/mvzt.gov.si/page uploads/pdf/informacijska_druzba/si2010.pdf

[6] State Portal for entrepreneurs, http://evem.gov.si/sp

[7] E-taxes Portal: http://eDavki.durs.si

[8] Reduction of administrative burdens; Programme of Measures for Reduction of Administrative Burdens, http://www.mju.gov.si/en/better_regulation/reduction_of_administrative_burdens/

[9] The User Challenge Benchmarking The Supply Of Online Public Services, http://ec.europa.eu/-information_society/eeurope/i2010/docs/benchmarking/egov_benchmark_2007.pdf

[10] Presidency Conclusions of the Brussels European Council (8/9 March 2007), http://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/ec/93135.pdf

Information Society and eGovernment Developments in Croatia

Božidar Kliček
University of Zagreb, Faculty of Organization and Informatics
Pavlinska 2, HR-42000 Varaždin, Croatia
E-mail: bozidar.klicek@foi.hr

Dijana Plantak Vukovac
University of Zagreb, Faculty of Organization and Informatics
Pavlinska 2, HR-42000 Varaždin, Croatia
E-mail: dijana.plantak@foi.hr

Keywords: Croatia, eGovernment, e-Croatia 2007 programme

Received: June 3, 2007

The aim of this paper is to provide insights into the current state in the field of eGovernment in the Republic of Croatia. In the paper a brief historical outline of the efforts undertaken so far is presented, certain key problems identified and development potentials in this area proposed. First, an overview of Croatian governmental organisations and actors that play an important role in adopting the eGovernment strategy is given. Then, results of the evaluation of the progress in the field of eGovernment are briefly presented by means of the European Commission's indicator for online availability, which lists twenty basic public services to be benchmarked. Next, the e-Croatia 2007 Programme *projects and areas for development are discussed, with the analysis of the achievements so far. Finally, key issues concerning eGovernment are summed up in the conclusion, along with the propositions for future work.*

Povzetek: Predstavljena je strategija razvoja elektronske uprave na Hrvaškem.

1 Introduction

In the 1980s Croatia was by far one of the most promising communist countries in Europe, and one of the countries in the region to claim the level of economic and democratic development not much unlike that in Western European countries. The disintegration of Yugoslavia as a multinational country was followed by an escalation of war and intensification of social, economic and political problems, all of which would provide a major hindrance to transition. This explains why Croatia is among the last post-communist countries to join the European Union. On the other hand, Croatia is commonly referred to as a stability factor, as a country at the forefront of the democracy movement in the Western Balkans, that is, ex-Yugoslav countries to have arisen from its disintegration. The last urgent conditions to be fulfilled by Croatia in the pre-accession phase, as defined by the European Commission, encompass jurisdiction, fight against corruption, and protection of ethnic minority rights.

One of the technologies which can lead to considerable improvements in all these components is eGovernment, which is why special emphasis should be put on it: not only can it increase the quality of life for its citizens, but it also contributes to faster integration of the Croatian society as a whole into that of contemporary Europe. Unfortunately, in its relatively brief history, eGovernment has been heavily marked by trends plaguing all developments in Croatia. Thus the state of eGovernment sophistication in Croatia is behind that of European Union countries, but ahead of Western Balkans countries. The Croatian Government, aware of the significance of this technology, has determined its priorities and trends with a view to making eGovernment one of the most propulsive segments of the Croatian state administration upon its eventual accession to the European Union.

2 History

2.1 eGovernment history in Croatia

In middle 70s, state administration and large companies applied information and telecommunication technologies (ICT) to a considerable extent, while in mid-80s ICT entered into small and mediumsized enterprises. However, in early 1990s, due to war, stagnation and even regression in the ICT sector was evident [3], leaving information society development and eGovernment strategies aside.

The Stabilisation and Association Agreement between Croatia and the European Union signed in 2001 [11] was the first step toward establishing stronger relationships with EU countries. It also emphased importance and need to further develop the information society by creating

legal and institutional frameworks as well as adopting standardized benchmark methodology to monitor the progress of the development.

In January 2002, the Croatian Parliament adopted a strategy entitled *Information and Communication Technology – Croatia in the 21st Century* and introduced the *General Measures for the Development of the Information Society*. At the end of year 2003, the Government of the Republic of Croatia adopted *e-Croatia 2007 Programme* following the guidelines of *The Action Plan e-Europe 2005*, supplemented with *The recommendations for the Action plan after 2005* (eGovernment beyond 2005) and finally harmonized with *The i2010 Initiative*.

The *e-Croatia 2007 Programme* goal is to "provide the Croatian citizens and the economy with the highest level of information services and the most widespread use and exchange of information, thus creating opportunity for their active participation in global developments" [1]. The programme includes the mechanisms of implementation and monitoring of the activities and projects planned for the periods of one year duration. The latest programme is elaborated in the *Operational Plan for the Implementation of the e-Croatia 2007 Programme for the year 2007* [14] and was published by the *Central State Administrative Office for e-Croatia*. It relies on previous *e-Croatia 2007 Programme for the year 2006* which resulted in significant improvements, especially in areas of the informatization of education (e-Education), the informatization of justice (e-Justice) and the informatization of health (e-Health) [4].

In the scope of eEurope, evaluation of information society development rate is conducted in sequential time intervals as standardized methodology. The results of evaluation are expressed with adequate indicators. In the field of eGovernment, the European Commission and member states have determined 12 basic services for citizens and 8 basic services for businesses, whose online availability is monitored [2]. Adopting the same methodology, three benchmark studies have been conducted in Croatia, which indicated distinctive progress in online availability of public services during last three years, 43.12% in services for citizens and 55.12% in business services (see Figure 1). This remarkable increase has enabled Croatia to take the leading position among other Western Balkans non-EU countries (see Figure 2) but when comparing the availability of public services on the Internet in Croatia and EU countries, a significant gap still exists (see Figure 3) [15]. Thus, more efforts are needed to adapt the Croatian legal and institutional frameworks in order to transform the Croatian society into an information society. Familiarization with the best practices from other countries, especially with those having a similar jurisdiction system, like Austria and Slovenia, may prove extremely helpful in that respect.

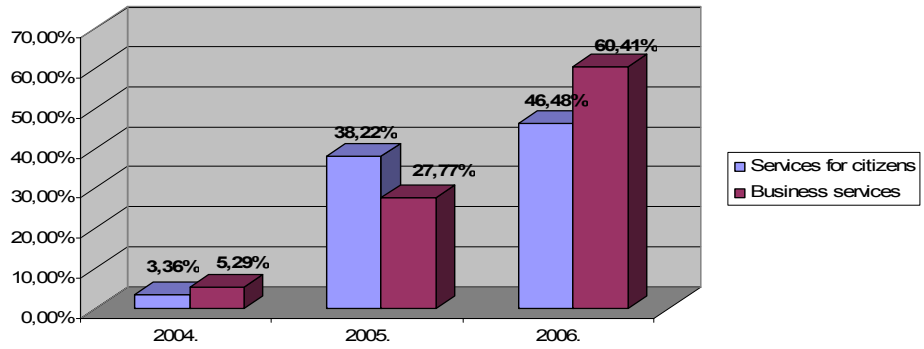


Figure 1: Progress of online availability of public services in Croatia from 2004 to 2006 (source adapted from [15])

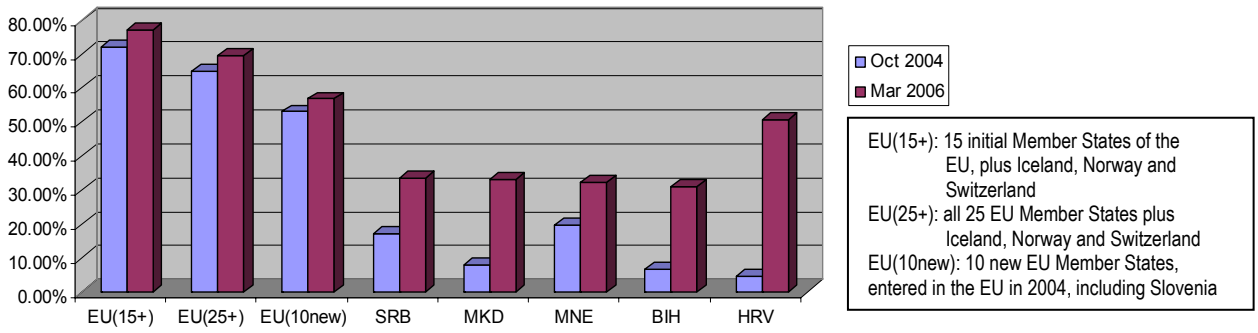


Figure 2: eGovernment benchmark of EU countries and Western Balkan non-EU countries (source adapted from [6])

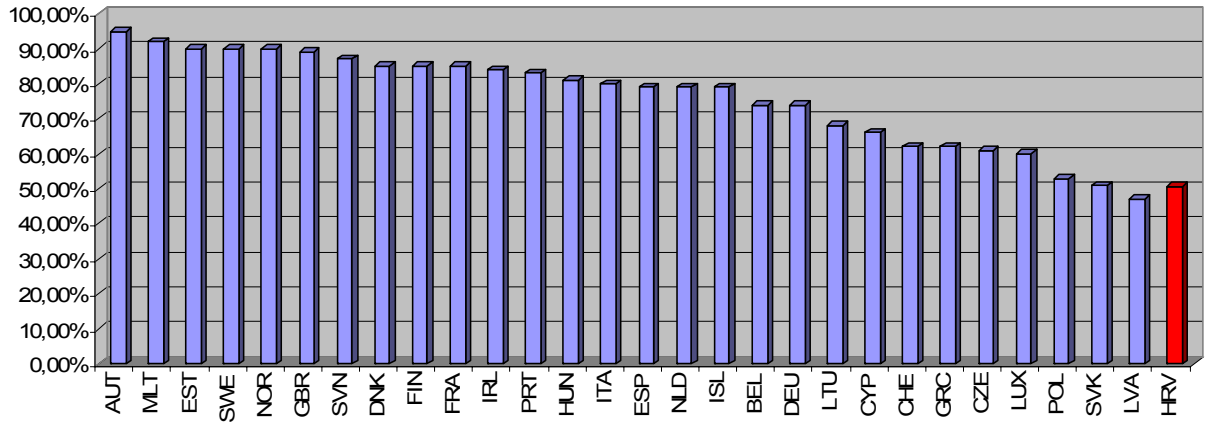


Figure 3: Online sophistication of public services in the different countries in 2006 (source adapted from [15])

3 State of the art

3.1 Croatian eGovernment strategy

For the purpose of providing proposals, opinions and expert views on different issues important for the state development, the Government of the Republic of Croatia forms its permanent working bodies: *Ministries, Offices of the Government, Central State Administrative Offices, State Administrative Organizations* and *Public Sector*.

Unit in place to design eGovernment policies and to coordinate ministries, regional and local governments in this sense is the *Central State Administrative Office for e-*Croatia of the Government of the Republic of Croatia. It is headed by the State Secretary and the two Deputy State Secretaries who are appointed to the Office. Internal structure of the *Central State Administrative Office for e-Croatia* comprises the following departments with a broad scope of activities:

- The Department for the Rationalization of Investments in Information and Communication technologies,
- The Department for the Coordination of the Implementation of the e-Croatia programme,
- The Department of International Cooperation.

The ultimate goal of the Office is to coordinate and implement activities of the e-Croatia 2007 programme to make quicker steps towards the information society following the recommendations of European Union and Lisbon Agenda [13].

3.2 e-Croatia 2007 Programme overview

The implementation of the e-Croatia 2007 programme is incorporated in the *Operational Plan for the Implementation of the e-Croatia 2007 Programme*. So far, three documents were issued: for the year 2004 and 2005, for the year 2006 and for the year 2007. All the bodies of state administration as well as other institutions, whose activities include the implementation of the e-Croatia 2007 programme, create institutional framework and participate in the programme preparation. The Government of the Republic of Croatia approves the annual operational plan, after it has been considered and evaluated by the *National Council for Information Society*. The bodies of state administration in charge of the implementation of particular measures submit their semi-annual reports to the *Central State Office for e-Croatia* regarding the current state of the implementation of those measures. The *Central State Office for e-Croatia* generates their consolidated report to be submitted to the Government of the Republic of Croatia and the *National Council for Information Society* [14].

The e-Croatia 2007 Programme is divided into two main areas:

- infrastructure, which comprises *Broadband, Interoperability, Information security*, and *HITRONet Network*;
- projects, which include *eGovernment (e-uprava), e-Justice (e-Pravosuđe), e-Business (e-Poslovanje), e-Education (e-Obrazovanje)* and *e-Health (e-Zdravstvo)*.

To realize both areas, besides an institutional framework, legislation is required, and in that respect Croatia has already adopted several acts to support the ICT development and its application in private and business sector, such as Financial Agency Act, Electronic Signature Act, e-Commerce Act, Registry Number Act, Act on Personal Data Protection, Telecommunications Act, etc.

3.3 e-Croatia 2007 Programme components

Broadband

A basic infrastructure of the information society and knowledge-based economy as its economical counterpart is high-speed Internet access. It must be available, affordable and useful to every citizen, household, school, company and public administration. In October 2006, the

Government of the Republic of Croatia adopted the *Strategy for the Development of Broadband Internet Access by the year 2008* and *Implementation Action Plan of that Strategy for the year 2007*. Strategy and Action Plan for the development of broadband Internet access create prerequisites for the accelerated development and adoption of this technology. The goal of the Strategy is the reduction of the gap between Croatia and EU countries as well as to encourage an even-paced development through all Croatian regions and population categories. Thus the Government of the Republic of Croatia decided to stimulate with about 5 mil euros the development of infrastructure for broadband in areas of the country where the interest of telecommunication operators continues to remain low, such as highland areas and islands [1].

Interoperability

Interoperability is defined as the ability of information and communication systems and business processes to support data flow and enable the exchange of information and knowledge. It can be achieved by adopting national and international technical norms [1].

In Croatia the *Croatian Standards Institute* is responsible to develop and align its standards to recommendations of international organizations for standardization. So far, some open standards are accepted to be implemented, such as Web Content Accessibility Guidelines 1.0 (implemented on all public administration web sites in 2007) developed and maintained by World Wide Web Consortium (W3C), standards for e-business developed and maintained by the Organization for Advancement of Structured Information Standards (OASIS), etc. Croatia is also started to develop an interoperability framework for several segments of public administration (The Customs Administration, The Tax Administration, Spatial Data national infrastructure) to join national information systems (IS) with ISs of EU countries [14].

Information security

The precondition for eGovernment implementation is data interchange, so a security policy should be created and information security standards applied in that environment. Therefore, Croatia adopted *the National Programme for Information Security in the Republic of Croatia* as well as *the Plan for the Implementation of the National Programme for Information Security in the Republic of Croatia for the year 2005*. The *National Programme* defines goals for information security at the national level, the jurisdiction and duties of particular institutions in the area of information security, as well as the necessary mutual coordination of all factors of information security [1].

In the year 2007 the focus is on the strengthening of the appropriate security norms concerning the implementation of information and communication technologies. According to this, two ISO norms were adopted: 'Information technology – Security techniques – Code and practice for information security management (HRN ISO/IEC 17799:2006)' and 'Information technology – Security techniques – Information security management systems – Requirements (HRN ISO/IEC 27001:2006)' [14].

HITRONet Network

HITRONet Network is an information-communication network for state administration interlinked into a unique communication infrastructure. It links various bodies of state administration and provides common Internet access, access by remote users, assistance to users, system security, overseeing and managing the system as well as the subsystem of common network and application services. The network is developing in phases, and the third phase anticipates broadening of the HITRONet access points to the level of cities, which will create conditions for a universal linking of state administration bodies into their virtual private networks [1]. It is projected that the end of 2007 will have comprised 80% of eGovernment services of the central state administration bodies within HITRONet [14].

eGovernment (e-uprava)

To enable the development of common electronic services and central access to information resources of the government administration, many projects have been initiated in the area of eGovernment. The main project that incorporates several electronic services is the HITRO.HR service intended for quick communication of citizens and business subjects with the state administration. HITRO.HR is based on the concept of "one-stop-shop" offering citizens and business subjects all the information about the required documentation, as well as the forms and money orders, at the HITRO.HR web site and HITRO.HR counters [7].

HITRO.HR comprises of several services: the establishment of Limited Liability Company, e-Regos (The Central Registry of Insured Persons**)**, e-Tax, e-VAT, e-Pension, e-Craft, e-Cadastre, and e-Corner. Majority of the services mentioned require authorization for accessing the service and authentication of forms by applying smart cards with a digital certificate issued by the Financial Agency.

Other projects that are taking place in the bodies of government administration are the e-Registries, the Electoral Register, the Central Database Registry on Personal Data and the Eurovoc Thesaurus.

e-Justice (e-Pravosuđe)

e-Justice projects were initiated to introduce information and communication technology into the judicial system. Projects serve to citizens, judges and other judiciaries by enabling access to legal databases and registers. The *e-Portal of the Ministry of Justice* provides information on the activities, reforms and functioning of the judicial system [1]. Some projects that are operational are: Integrated communication system for managing court cases (ICMS), the e-Land-registry Certificates Project,

the e-Cadastre Project, the e-Court registry Project, the e-Judicial Practice Database Project, the Judges Web etc.

e-Business (e-Poslovanje)
The goal of e-Business projects is creating favourable conditions for the successful development of electronic business. Legal framework with several laws and decrees already exists, and the Strategy for the development of e-business should be adopted by the end of the year 2007. Two e-Business projects are incorporated into the HITRO.HR service (the e-REGOS project and the e-VAT project). Other projects are the e-Crew project and the e-Customs Project.

However, despite the predominantly positive attitude towards electronic business, its potentials have not been utilized sufficiently. One study showed that main reasons for not using public services for business were lack of information about the services, habit of doing business in traditional way, the perception that a company doesn't have sufficient technical conditions for e-business and information security issues [5]. In order to promote its services for business, the *Ministry of Economy, Labour and Entrepreneurship* together with the *Croatian Chamber of Economy*, initiated the campaign to ensure a broader acceptance of electronic business.

e-Education (e-Obrazovanje)
Owing to the scope of its vision, the e-Education project is probably one of the most ambitious projects among e-Croatia projects. It includes several aspects, such as information-service infrastructure, application of ICT in teaching, and the development of standards and application of certifications in education, necessary for Croatia's further growth as a knowledge country [1]. Some of the projects under e-Education umbrella are the following: the Information System of Elementary and Secondary Education, the Information System of Higher Education Institutions (ISVU), GigaCARNet, Mobile CARNet, free broadband access in student dormitories (StuDOM), the Croatian National Educational Standard (HNOS), the ECDL (European Computer Driving Licence) project, e-Indeks (student's electronic card), etc.

e-Health (e-Zdravstvo)
The computerisation and integration of the national health care system is one of the strategic goals of the Government of the Republic of Croatia, which is to provide better service to patients, improve quality of life and decrease the health care system's costs. An integrated computer system for primary health care has already been released in 350 medicine practices and has brought several advantages: doctors have access to patients' medical records, laboratory test results are forwarded to patients' doctors and patients can schedule appointments via Internet.

By the end of 2008, all primary health care institutions will have been connected to the integrated computer system [1]. The informatization of the hospital information system (IBIS) as well as that of the system of social welfare is also in progress. The HZZO Portal (the portal of the Croatian Institute for Health Insurance, http://www.hzzo-net.hr/) now offers a possibility to fill in an electronic health insurance form.

4 Conclusions and future work

Croatia has realised its national ICT strategies and incorporated eGovernment into its National Strategy by introducing the *e-Croatia 2007 Programme* in the year 2003 and annually upgrading operational plans and activities within the Programme ever since (the latest version of the document being issued in the year 2007). By implementing the *e-Croatia 2007 Programme*, the Republic of Croatia will meet the preconditions for increasing the competitiveness of the country a whole as well as enable its equal and active participation in the development of the knowledge-based society, a much sought-after goal.

The progress in online availability of public services achieved so far, measured by the standardized EU methodology, shows that Croatia have already made significant steps in the implementation of eGovernment. The changes that have been occurring cannot only be monitored by means of standard EU instruments for measuring annual progress but are also evident in everyday life. In order to popularize eGovernment services, the *Central Administrative State Office for e-Croatia* initiated a new HITRO.hr service called "e-KUTAK" (*e-Corner*), a free educational service intended to encourage entrepreneurs and craftsmen to enter in the world of electronic business. National and commercial television channels have been broadcasting video clips promoting governmental e-services for both citizens and entrepreneurs (HITRO.hr and the Croatian Government portal, http://www.vlada.hr). Very recently, a new internet portal of the Government of the Republic of Croatia, called "Moja Uprava" (*My Government*, http://mojauprava.hr/ [10]), has been implemented. Based on the experiences of the other countries, the portal offers relevant administrative content and governmental services to citizens, businesses and other users, all in one place. The portal has been developed according to guidelines of W3C, making it accessible and usable to a wider public.

However, 42.6% of state, local and public sector offices and service providers still do not have a web site [15]. Many existing web sites, on the other hand, are not compliant with WCAG standard and are not multilingual. Also, some of the e-Croatia 2007 Programme strategic documents are only available in Croatian. New annual report regarding the current state of the implementation of the activities and measures set up in the *Operational Plan for the Implementation of the e-Croatia 2007 Programme for the year 2007* should be issued by the end of the year to show what have been achieved in 2007.

According to the '2006-2007 Global Report on Information Technology', published by the *World Economic Forum*, Croatia still falls behind when the usage of IT in public state administration processes and services is concerned. In Croatia, the *National Competitiveness Council* has therefore emphasised the importance of development of eGovernment in general, at both the national and local level, as one of the major recommendations for increasing its ICT competitiveness, which primarily refers to developing the systems of e-Procurement, e-Health and e-City. Other important recommendations include the necessity of enabling broadband Internet access to all citizens and promoting e-business among business entities [12].

In the next period Croatia should recognize which EU solutions to adapt and which issues to develop by its own means and efforts. One of the priorities introduced in both *The i2010 Initiative* [8] and the *Operational Plan for the Implementation of the e-Croatia 2007 Programme for year 2007* [14] is the innovation and investment in ICT research. Croatian scientists are already participating in the development of the *European Research Area* (ERA), by joining a whole range of FP5, FP6 and FP7 projects. The knowledge and experience arisen from conducting research projects should be aimed on creating and upgrading the curricula in academic institutions. Considering eGovernment issues, academic institutions in Croatia have already introduced the training of civil servants (through training courses). The next step forward should be the establishment of a postgraduate specialized study in eGovernment, based on examples of the best EU practices, to provide knowledge for a successful eGovernment projects delivery and eGovernment systems management.

At this stage the cooperation with other countries to have arisen from disintegration of the post-communist Yugoslavia is vital, both when providing examples of best practice in solving state administration problems and, above all, when ensuring technological interoperability essential for establishing a quality relationship with future EU member countries coming from the region, are concerned.

References

[1] Central State Administrative Office for e-Croatia: *e-Hrvatska*, http://www.e-hrvatska.hr/, *e-Croatia,* http://www.e-croatia.hr/

[2] Cap Gemini: *Web-based Survey on Electronic Public Services, Results of the first measurement,* October 2001. Retrieved June 20, 2006, from http://ec.europa.eu/information_society/eeurope/2002/documents/CGEY%201st%20measurement.pdf

[3] Croatian Chamber of Economy, Information System and Statistics Department: *Croatian ICT Sector*. Zagreb, July 2004. Retrieved June 20, 2006, from http://www2.hgk.hr/en/depts/IT/hrv_ict_sektor.pdf

[4] e-Croatia: *Government adopted resolution plan for e-Croatia for 2007*. Zagreb, May 10, 2007. Retrieved May 13, 2007, from http://www.e-croatia.hr/

[5] Gfk Group: *Public services on the Internet – obstacles for utilization and priorities for informatization*. September 2006. Retrieved May 13, 2007, from http://www.e-hrvatska.hr/repozitorij/dokumenti/downloads/Istrazivanje_prioriteta_i_prepreka_u_koristenju_javnih_e-usluga.pdf

[6] Gusev, Armenski (April 2006): *Gap Analysis of eGovernment in Western Balkans*. Retrieved June 20, 2006, from http://www.metamorphosis.org.mk/

[7] *HITRO.HR* - service of the Government of Republic of Croatia intended for quick communication of citizens and business subjects with the state administration, http://www.hitro.hr

[8] *i2010 – A European Information Society for growth and employment.* COM(2005) 229 final. Brussels: Commission of the European Communities, 1.6.2005. Retrieved November 12, 2007, from http://eur-lex.europa.eu/LexUriServ/site/en/com/2005/com2005_0229en01.pdf

[9] *My Government*, a portal of the government of the Republic of Croatia, http://mojauprava.hr/

[10] Ministry of Foreign Affairs and European Integration: *Stabilisation and Association Agreement (in Croatian)*, Retrieved November 12, 2007, http://www.mvpei.hr/ei/default.asp?ru=143

[11] National Competitiveness Council: *Recommendations for increasing ICT and technological competitiveness of Croatia (in Croatian).* April 2007. Retrieved May 13, 2007, from http://nvk.multilink.hr/dokumenti/119___Preporuke%20za%20povecanje%20ICT%20konkurentnosti%20Hrvatske.pdf

[12] *Operational plan for the implementation of e-Croatia 2007 program for 2006* (only in Croatian). Zagreb: The Government of the Republic of Croatia, The Central State Administrative Office for e-Croatia, April 2006. Retrieved April 16, 2007, from: http://www.e-croatia.hr/repozitorij/dokumenti/downloads/plan_provedbe_e-Hrvatska_za_2006.pdf

[13] *Operational plan for the implementation of e-Croatia 2007 programme for 2007* (only in Croatian). Zagreb: The Government of the Republic of Croatia, The Central State Administrative Office for e-Croatia, May 2007. Retrieved May 13, 2007, from http://www.e-hrvatska.hr/repozitorij/dokumenti/downloads/Operativni_plan_eHrvatska_za_2007.pdf

[14] T&MC, e-Hrvatska: *Benchmarking study on online availability of public services - third research for year 2006* (presentation for the press, *in Croatian)*, Retrieved July 13, 2007, from www.e-hrvatska.hr/

eGovernment Services in Bosnia and Herzegovina

Senid Gerin and Brane Vujičić
B@IT, Asocijacija za informacione tehnologije u BiH
Marka Marulića 2/III, 71000 Sarajevo, BiH
E-mail: info@bait.org.ba
Web: www.bait.org.ba

Keywords: eGovernment, ICT, Bosnia and Herzegovina

Received: August 28, 2007

During war years (1992-1995) in Bosnia and Herzegovina development of information society and usage of information and communication technologies (ICT) were halted whilst expansion of the ICT recorded the highest figures in the whole world. As result, Bosnia and Herzegovina (B&H) society is at bottom line regarding the development and usage of IC technologies.

Post-war development of the information society has not been following world trends because of the complexity of B&H constitution that limits development on many B&H segments including the ICT.

Basic prerequisites for further development of ICT have been established through state-level adoption of documents covering the Policy, Strategy and Action plan for the development of the information society in B&H for the period 2004 – 2010.

According to several analysis reports on this sector published by WEF, UNDP, the Secretariat of eSEE, telecomm operators and other institutions, wherein world standards have been taken as reference, the basic ICT infrastructure in B&H is rather well developed what makes suitable environment for development.

In the post-war development of the B&H information society the international community has provided substantial support through its' organizations including UNDP, USAID, EC and others.

Through increased engagements of all relevant players, through adopted legal framework and through the framework to be adopted substantial improvement in the development of the B&H ICT can be expected.

Povzetek: Članek podaja pregled storitev e-uprave v Bosni in Hercegovini.

1 Introduction

In order not to remain aside of the global changes, which include establishment of a contemporary information society that is based on intensive application of knowledge and information as well as usage of IC technologies in everyday life, and respecting experience of other countries, Bosnia and Herzegovina has adopted policy and strategy that are going to be the basis for development of information society as key guidelines.

The policy of B&H information society development represents basic- and framework document that is going to be a basis for creation of new legislation, rules and other legislation acts as well as the basis for making of further decisions on development direction, action plans and priorities at the levels of B&H state and entities.

The policy defines strategy of information society development regarding specific development areas such as ICT industry, eBusiness, eHealth, eGovernance, eLegislation, information society and sustainable development.

In November 2005 the B&H Council of Ministers adopted the Policy, Strategy and Action plan of the information society that is composed of five development pillars for the Legal infrastructure, eEducation, eGovernance, ICT infrastructure and ICT industry. The strategy covers 2004 – 2010 period and represents conceptualization of needs and commitments expressed in the Policy of the B&H information society development, as well as the methods to achieve them. The strategy offers understanding of the methods to be used so as to accomplish commitments defined in the B&H information society development strategy.

Each of the five development pillars is shown in the Strategy as well as a brief description of the current state, basic strategic guidelines to achieve the vision and

specification of concrete actions to be taken respecting the strategic guidelines.

The Action plan has been adopted whereby the plan describes specified actions as well as progress tracking indicators.

The achieving of goals as defined in the Policy, Strategy and Action plan does not record dynamics desired.

The basic problems are reflection of complicated composition of the B&H state that is composed of two entities and district area. In the state there is a three level structure with 14 parliaments: one at state level, two at entity level, one at district level and 10 parliaments at canton level. At various level individual issues get solved and a legislation procedure has to be carried out at all levels. Such approach mostly slows down establishing of the legal environment for efficient realisation of the Strategy for B&H information society. An example of this kind of obstacles the Act on the agency for the B&H information society can be mentioned because though it has been in procedure for adoption for three years the final proposal of the Act still has not been brought up yet. Basic tasks of the Agency are: strategic planning, standardization and coordination, and organizing of the environment to realise adopted Policy, Strategy and Action plan for the B&H information society development in the 2004 – 2010 period.

2 B&H Legal framework for information society

According to constitution the B&H state has the legislative power bodies and the Parliament embodied in the Council of ministries that has governmental competences. The state is composed of two entities where one of them is the Federation of Bosnia and Herzegovina (B&H F) entity composed of ten cantons where each canton has its parliament and government. The other entity is the Republika Srpska (RS) that has its' Parliament and Government. In the state there is the Brčko District that has it's Parliament and Government.

Existing competencies of the B&H Council of Ministries are set as a framework; however, they are sufficiently broad basis for decisions making and work of the B&H common institutions. Moreover, the B&H common institutions competencies covering various social and business areas and expected to broadened.

It can be justified assumption that in a foreseeable time the cooperation of the Council of Ministries, law enforcement- and legislative power bodies of entities, the B&H F, RS and Brčko District, and vice versa. Besides others, the requests of the environment are directed towards more efficient functioning of institutions at all levels throughout the state. One of the basic prerequisites to suite this request is that number of decision making levels should be reduced, both in terms of the legislative power and in terms of the governance.

In terms of geography, economy and industry, in terms of the population density and migration, Bosnia & Herzegovina cannot afford existence of legislative power at the levels of the state and its' entities, district and ten cantons. The aforementioned has resulted with the existence of the legislations and rules at levels of cantons, entities, state level as well as the existence of the "roof" legislation that the basis for the entity legislation covering individual areas.

Overcoming of this problem is prerequisite to access regional and other integrations; the only way to overcome these problems is strengthening of the legislative power at the state level. This is particularly important for the area of usage of IC technologies because their usage does not allow barriers of this nature and of these kinds.

Bringing up of the state level rules became important in 2002 when many state level acts, which are important for these matters, were enacted. It is particularly worth mentioning a set of positive rules, which cover ICT issues related to the need for introducing basic registries and which are related to the application of the ICTs in various work and living areas that encompass some of 20 basic G2C and G2B public services.

Existing state of the eLegislation in the legal system of Bosnia and Herzegovina

Legal framework in B&H legal system is currently not developed and it is not sufficient for employing of the ICTs on a large scale. Establishment of adequate legal framework at the state level for the eLegislation is the basic prerequisite to achieve the goal called information society entering.

Particular themes could be acts and rules that request application of the ICT as an indispensable prerequisite, where the act enforcement presumes that that the ICT is used in governance bodies, both in state and in companies. Enacting rules and regulations and getting them enforced is a slow process. Additionally, some acts are not harmonized: for instance the Act on public listed companies of B&H F defines that the shares are to be issued in electronic form and that they have to be registered at the Registry of bonds and shares. However, this has not been directly implemented in the Act on the registry of bonds and shares. This is an illustration of the need for compulsory usage of the ICT for registering of data and tracking of changes of individual bonds and shares. There is a practice that the act enforcement gets multiply postponed because of lack of prerequisites for their application.

The legal frame for eLegislation does not exist at any level of B&H state organization. There are just acts and regulations for some areas (for instance banking) what cannot be considered as sufficient in any way. The aspects of the ICT usage are just one segment of the eLegislation. eLegislation building is the building of a special pillar for the needs of the information society. The eLegislation should be a means for building of consistent legal framework at the state level that is going to be a basis for the implementation of all five pillars what should enable joining of the B&H to the information society.

3 eGovernance

Generally considered the eGovernance describes processes being performed between governments at one- and citizens and businesses at other side even in case when performed partially using the ICT. The term "eGovernance" is understood as a communication between the citizens and government institutions. This kind of communications is presented with abbreviation G2C (government to citizen). eGovernment is used for communication between business and government a well known acronym G2B, The processes can be performed through the Internet, dial-up phone or any other ICT infrastructure that is suitable for communication. Regarding citizens one of the requests on the eGovernance systems is to make the processes more efficient by using the ICT.

Reform of Public Governance institutions and introduction and employing of state-of-art technologies is one of priorities within the building and functioning of the B&H governance system. Respecting the economy and political environment changes and respecting the reorganization of the whole structure of the governance and management, the basic intention of this reform is to determine essentially newly positioned organization and tasks of the public governance institutions so as to ensure faster and less expensive, and better services to the citizens. The reorganization of the managing system structure should be performed so that it provides support to the expected changes that provide more viable and more contemporary vision of the B&H future.

It is important to stress the difference between the terms eGovernment and eGovernance. The term eGovernment is more related to technical solution providing more efficient and transparent operation of the government bodies by using ICT whereas the term eGovernance is related to overall process often including reengineering of the business processes as well as a reform that is more deeper than just introducing of the IC technologies in the work of the governance bodies and government. The transition process and introducing of the eGovernment in B&H should be considered as an evolutionary process in which there are periods of fast changes and standstills that are necessary to understand the changes and to be implemented.

Should Bosnia and Herzegovina be associated to the European Union by 2010 it will have to carry out reforms in public governance institutions and will have to switch to the eGovernance concept.

Public administration provides services to citizens in 146 municipalities, 2 entities, in Brčko District and at the level of B&H state. At each level there are problems with administration work, such as follows:

- The acts are adopted by 14 parliaments (in some segments conflicting),
- Administration does not operate according to European standards,
- Principles and practice at the level of an municipality or a canton are different from principles adopted in other municipality, canton or B&H F,
- Lack of transparency regarding to work and finance matters,
- Tendency to unify and to connect databases is poorly pronounced (some projects at the level of B&H, for instance CIPS, DGS, Customs and others are exceptions),
- There is neither horizontal nor vertical electronic communication,
- Various operating systems, application and database systems because no hardware and software policy and standards have been adopted,
- There is no global plan for introducing information technologies in the state administration,
- Existing system operate autonomously - as islands so that they cannot provide citizens with information,
- Partly outdated equipment does not enable suitable networking,
- Usage of state of art communication (such as video conferencing, email etc.) is at minimum
- Web-pages of the public administration provide citizens with not many services, only not frequently updated information, they rarely include forms that can be printed locally,

In essence the eGovernance concept does not function.

According to Europe 2002 indicators there are four phases of the development of eGovernance services:

Phase 1 – information. One-way services to citizen, such as reports and on-line documents.

Phase 2 – one-way interaction. Web portals offering possibility to print forms and other documents directly from the web. Usually there is an option for so-called download forms.

Phase 3 – Two-way interaction. Public web portal enables citizens to electronically submit request to the Government by using web or e-mail.

Phase 4 – complete electronic data processing. Complete communication with the government officials goes through the ICT. The request tracking system operates electronically and there is no exchange of paper form documents throughout the process.

Besides the aforementioned "0 phase" has been introduced. It describes absence of any form of the eGovernance. This phase includes eGovernance having portal but its' content does not provide any relevant information.

According to the World Economic Forum classification of the government on-line services, Bosnia got 2,12 point mark (maximum is 7, Singapore) ranking it as 71st of 104 countries.

By now in Bosnia and Herzegovina there has not been a study of this kind; however, there are indications that most government institutions are in phase 0, and only few are in phase 1 of the eGovernance development. The Sector Status Report from 2004 reads that only 61% of the government institutions have access to the Internet. With support of international organizations, in the last several years some portals offering more than just static information have been developed, for instance the register of vehicles.

	Public Services for Citizens
1.	Income taxes: declaration, notification of assessment
2.	Job search services by labor offices
3.	Social security contributions: a) Unemployment benefits, b) Family allowances, c) Medical costs (reimbursement or direct settlement), d) Student grants
4.	Personal documents: a) passport, b) driver's license
5.	Car registration (new, used and imported cars)
6.	Application for building permission
7.	Declaration to the police (e.g. in case of theft)
8.	Public libraries (availability of catalogues, search tools)
9.	Certificates, request and delivery: a) birth, b) marriage
10.	Enrolment in higher education / university
11.	Announcement of moving (change of address)
12.	Health related services (e.g. interactive advice on the availability of services in different hospitals; appointments for hospitals.)
	Public Services for Businesses
13.	Social contribution for employees
14.	Corporation tax: declaration, notification
15.	VAT: declaration, notification
16.	Registration of a new company
17.	Submission of data to statistical offices
18.	Customs declarations
19.	Environment-related permits (incl. reporting)
20.	Public procurement

According to suggestions of the eEurope and eEurope 2005 standards development of eGovernance is evaluated through development of 20 public services (see table above).

By now there has been no official analysis of eGovernance state for these services; however, it can be concluded that there is no service offered completely through the eGovernance. The Agency for Public Administration partially offers job search on-line through portal. The ongoing projects carried out for agencies for employment in entities will offer job-search services and processing of unemployment compensation whereby both services are going to be fully integrated and will be categorized as phase 3 of the eGovernance. National University library develops integrated overview of individual library funds.

Respecting the B&H information strategy development some activities are being carried out to solve C2G and B2G requests for many services. At this moment (2007) the realizations of individual services are at different level of implementation: from initiation-, mobilization-, development, over implementation to project exploitation phase. Most of the authorized institutions have their web sites with basic information for citizens and other parties interested in, as well as for publishing of rules and regulations related to individual areas.

Because of government specific structure and because of specific state organization, so as to provide these 20 services it will be needed to adjust solutions to suite specific environment. The experiences in the region can be useful but cannot be cloned in B&H what requires an additional effort to realise specific public services.

4 Conclusion and future

The vision of the new state of eLegislation is coherent employment of project solutions because the current state of eLegislation is such that eLegislation scope is neither defined nor regulated through legislation. Coherent employing application of project solutions represents the vision of the new state of eLegislation. Enacting of necessary rules and regulations, as well as amendments to the acts already enacted is one of prerequisites for a new state of eLegislation. Adequate monitoring, user training, setting up of adequate organizational eGovernance bodies at the levels of state and eventually at the level of entities are additional prerequisites. Establishment of a separate body at the state level will be indispensable if project defined goals are to be achieved.

The B&H legal system changes that are going to be indispensable are as follows:

- Defining of law of legislation at the state level, and at entity levels and harmonizing them with the state level legislation.
- Reduction of decision making levels,
- Harmonization of the current legislation rules with the new eLegislation,
- Setting up of organization for education, monitoring and all other post project activities.

Which steps are necessary for new state of eLegislation?

In the first phase this should be completion of existing project activities related to the strategy and action plan as follows:

- ICT infrastructure,
- ICT industry,

- eEducation,
- eGovernment and
- eLegislation

as a legal framework for previously mentioned project activities. All these steps are going to set up the basis for information society entering in a foreseeable time period. Essentially, they can be divided in the following phases for the period of eLegislation project performance:

- Production of new legal rules and regulations,
- Monitoring and
- Education.

Post project period is going to be performed in phases as follows: implementation, continual monitoring and education.

Time milestones given in the Strategy and Action plan should be revised and modified in accordance to the real situation.

The second phase should coincide with the whole process cycle of project and with other activities related to information society entering.

The rules and legislation to be enacted and amended can be found in the Strategy for B&H information society development, eLegislation chapter.

Strategic directions and guidelines for work related to the eGovernance are shown in details covering specific areas of work and development identified in the Policy of the B&H information society development. The areas are as follows: re-engineering of the governance, technology-development basis, infrastructure, interoperability of organizational units, fundamental registries, security, common functions of the governance bodies, special functions per organizational units of the governance bodies, eDemocracy, eServices, portals and access points.

The goal is to identify all fundamental problems, key prerequisites and tasks for organized and systemic development of the eGovernance and to identify the projects that would enable the fastest realisation of the eGovernance concept. The following issues should be emphasized: fundamental (research and development) projects, project that would provide maximum return on investments, basic services as planned by the European Union, as well as other services and functions which would be useful if implemented at the first phase of implementation. For each area it is necessary to follow respective worldwide trends and experiences, goals to be achieved, as well as direction of work with necessary tasks so that the goals can be achieved. Besides identifying the technical and technological issues, and identifying the services that should be fully automated it is necessary to provide coordination of local- and higher levels activities whereby the authorization and autonomy of local level governance institutions has to be respected to maximum, and it should provide horizontal networking of organizational units (for instance association of the municipalities and cities in Bosnia and Herzegovina in the eGovernance Development Forum) and it should provide coordination and development of the introduction of eGovernance in each organization unit of governance institutions (governments, municipalities).

References

[1] Policy of Information Society Development in BiH, Council of Ministers of BiH nov-2004. http://www.is.gov.ba/downloads.php

[2] Strategy of Information Society Development in BiH, Council of Ministers of BiH nov-2004. http://www.is.gov.ba/downloads.php

[3] Policy of Information Society Development in BiH, Council of Ministers of BiH nov-2004. http://www.is.gov.ba/downloads.php

[4] Study on eReadiness for employment of eGovernance in rhe Northeast region of B&H - EXIT IT Business support centre http://www.itbusiness.ba/ITB/Documents/Brosure/eGovStudy.pdf

[5] eEurope 2002 Final Report, Commission of the European Communities, February 2002. http://ec.europa.eu/information_society/eeurope/2002/news_library/documents/acte_eEurope_2002_en.doc

[6] The Global Information Technology Report, World Economic Forum, 04/05. Issue 9.02 Government on-line services.

[7] eSEEurope Regional Information and Communication Technologies Sector, Status and Usage Report: Building an Information Society for All, UNDP and Stability Pact for South Eastern Europe, oktobar 2004.

[8] Online Availability of Public Services: How is Europe Progressing?, European Commission Directorate General for Information Society and Media, October 2004.

Survey of eGovernment Services in Serbia

Zoran Babović and Darko Jović
Innovation Centre of School of Electrical Engineering, University of Belgrade
Bulevar Kralja Aleksandra 73, 11120 Belgrade
E-mail: zbabovic@rti7020.etf.bg.ac.yu, sparrow@galeb.etf.bg.ac.yu, http://ic.etf.bg.ac.yu/

Veljko Milutinović
School of Electrical Engineering, University Of Belgrade
Bulevar Kralja Aleksandra 73, 11120 Belgrade
E-mail: vm@etf.bg.ac.yu, http://www.etf.bg.ac.yu/

Keywords: eGovernment, strategy, online services, sophistication, interoperability

Received: August 26, 2007

Serbia recognized the importance of the information society early in the 1989, when regulatory activities and organizational preparations in this area started. Despite of its early start, Serbia is now behind the European countries regarding the eGovernment services. The factors such as political situations, lack of capital, absence of appropriate legislation, etc. slowed Serbia in its development of the eGovernment services. This document describes the history of eGovernment in Serbia, gives overview of current state of eGovernment services, and presents strategy and actions plans for future development of interoperable eGovernment services.

Povzetek: Podan je pregled e-uprave v Srbiji.

1 Introduction

In this paper we tried to follow the structure of typical eGovernmnent report for EU countries done by EU IDABC eGovernment Observatory, but with little modifications.

Sections from 2 to 5 are common: we started with description of history of eGovernment services in Serbia, then with existing national strategy, legislative framework, and key organizations responsible for area of eGovernment development.

Sections 6 and 7 contain information about current implementation state. Infrastructure necessary for efficient eGovernmnet services is described in section 6, and in Section 7 is given rank of sophistication of standard online services for citizens and companies in Serbia.

There are initiatives (e-SEE Agenda) and EU projects (We-Go) that try to improve sophistication level and expertise of eGovernment services in region of Western Balkan countries. Therefore, for enabling better regional development we give a stress on description of current state in the governmental institutions responsible for services that tend to be cross-border transactional. Section 8 contains such information based on public available data.

2 History of eGovernment services in Serbia

Recognizing the importance and necessity of information society development, institutions of the Republic of Serbia, a bit earlier than other countries in the region, have started regulatory activities in this area [5] [6]. Since October 1989, in Serbia regulatory activities were started and organizational preparations on institutionalization and building of information system of governmental institutions. On the other hand, since 1996, realization of important projects was started in Serbia. However, great number of realized projects was not put in exploitation.

Here is the chronological survey of acts and projects in recent period.

Adopted acts:

- October 1989: Execute Council of Serbia Parliament adopted The Guidelines for building and adoption of the governmental institutions information system projects
- July 1990: Execute Council of Serbia Parliament adopted The Legal act on securing and protecting of governmental institutions information systems
- June 1992: Institute "Mihajlo Pupin, Belgrade, made the document: Conception of governmental institutions information system development. In 1995. this study document was adopted as an expertise base for creation of the

Law on Information system of Republic of Serbia
- March 1996: The National Assembly of the Republic of Serbia adopted the Law on Information System of the Republic of Serbia
- February 1997: Government of the Republic of Serbia, adopted the legal act on program of development of Information System of the Republic of Serbia, in order to enforce Law on Information System of the Republic of Serbia
- November 1997: Federal Government adopted the document: Strategy for further development of information technology in Federal Republic of Yugoslavia
- May 1998: Government of the Republic of Serbia adopted the Legal act on program of development of Information System of the Republic of Serbia
- May 1998: The Law on Personal Data Protection was adopted in 1998 on a Federal level
- April 2000: Government of the Republic of Serbia adopted The Legal Act on program of development and operability of information system of the Republic of Serbia in 2000.
- October 2001: Government of the Republic of Serbia adopted The Program of information system development of the Republic of Serbia
- October 2002: Minister for science, technology and development has signed Agenda for information society development in the South East Europe

Projects:
- July 1996: The Project of database for standards from information technology area was adopted. Database is implemented in Office for common operations of governmental institutions
- January 1997: The Project of database about codes and classifications in governmental institutions was adopted. Database was implemented in National Agency for Statistics.
- May 1997: The Project of transitional solution of common database of information system of the Republic of Serbia was adopted. According to the project, the common database contains data from Registry of citizens, Registry of legal entities, and Registry of land units, as well as classifications and codes for data usage. Conceptually, the Common Database represents core that connect and concatenate data from information subsystems databases.
- 1997: The Project of data dictionary of the information system of the Republic of Serbia was adopted. Among of three planned segments of the project, one was realized: Description of the organizational structure of the institutions, which was implemented in six institutions
- January 1998: The Project of common computer – telecommunication network of governmental institutions was adopted. In chime of the project, the main communication node of the network was equipped.
- February 1998: The Project analysis on reasonableness of software tools usage in information system of the Republic of Serbia was adopted.
- 1998/99: The Project - General Program system for administrative business was adopted. As well: Prototype of the general program system for automation of administrative business in governmental institutions and Prototype of program system for automation of administrative business of General Government Secretariat was done
- 2000/01: The Project of information system about personal status of citizens was adopted – Program system for computer data processing about personal status of citizens, which evolve three functional unit/subsystems: "Master register", "Electorate" and "Citizens"
- 2001: The Project of unique Register of streets and home numbers in the Republic of Serbia (Address Register) and Program system for maintaining of Registry were adopted.

Significant results in realization of important eGovernment projects were visible from 2002, when technological prerequisites were satisfied for intensive and efficient development.

3 Strategy for eGovernment development in Serbia

As a member of Initiative for Electronic South Eastern Europe (eSEE) which functions inside of the Stability Pact for South Eastern Europe, Serbia has signed international agreement "Agenda e-SEE for information society development" (Agenda e-SEE) on October 29th 2002., as a basic document for development of information society in this region. This agreement is in accordance with action plans e-Europe 2002 and 2005 and plan e-Europe+ for countries candidates and stands as confirmation of readiness of South Eastern European countries to work on development of information society in accordance with IT development processes in Europe [8].

The Minister of Telecommunication and Information Society Mrs. Aleksandra Smiljanic has ratified the new e-SEE Agenda+ on October 29th 2007. This new agenda represents the regional plan for activities in purpose of information society development for period from 2007 to 2011.

As defined in Agenda e-SEE, countries of South Eastern Europe will undertake concrete actions considering local constrained factors in the following areas:
- Adoption of work policies and strategy for information society

- Adoption and enforcement of legal infrastructure for information society in accordance with *acquis communitaire* established in EU countries
- Establishing regional collaboration and state mechanism for conduction and promotion of the information society development

In the Republic of Serbia, basic goals and directions of government strategy in the area of information society development and development of eGovernment is defined in the document "Strategy of Public Administration Reform in the Republic of Serbia" (November 2004) and this strategy represents framework for the Strategy for Information Society Development.

In this strategy is emphasized that as one of concepts and measure of reform success is inclusion of new information and communication technologies in public administration work.

By late 2004, local UNDP office has initiated work on National strategy for information society development.

Finally, Strategy for Information Society Development was adopted in October 2006 [4]. In this strategy, priorities and goals for information society development are set up, necessary institutional and legislation framework for such development is defined, strategy for establishing efficient national communication and information infrastructure is proposed, and besides special stress on eGovernment field, necessary steps and desired goals are stated in the area of e-Education, e-Health, e-Business, and e-Banking. In the special section development of business sector is described, which gives services and offers products from domain of ICT. Also, a number of measures are accepted for the sake of monitoring the successfulness of achieved goals of the strategy, as well as assessment of overall information society development. At the end of the strategy a one-year Action plan is given, with desired goals in certain segments, necessary activities, and deadlines for their accomplishment.

Special section in the SISD is related to eGovernment field. The main goals of eGovernment development are:

- Modernization of the public administration
- Development of the national economy
- More in-depth engagement and involvement of citizens in democratic processes

The concept of eGovernment anticipates interactive electronic services customized to the needs of citizens and businesses. Such services are integrated on all levels of public sector.

In order to realize such concept of eGovernment, the following principles of implementations are proposed in the strategy:

- Access for everybody (public services have to be available to all citizens)
- Security and protection of privacy
- Open system (eGovernment services should use solutions based on open standards and open formats)
- Coherency and functional integrity (eGovernment is coherent system, where integrity and common functionality of various heterogenic parts are reached through standardization and coordinated development)
- Autonomy in development (Every governmental institution or public organization should autonomy develop and manage their own subsystem according to previously agreed eGovernment standards and national plan of development)
- Flexible and modern ICT solutions
- Reliance on national ICT sector and academic/research community

It is stated that is necessary to develop the following key components of eGovernment:

- Electronic public services (Development of portal of all eGovernment services, services for citizens G2C, services for businesses G2B, services for administration G2G, internal electronic services among governmental institutions necessary for all other services)
- Telecommunication infrastructure
- E-Payment
- Security
- Standards
- Legislative infrastructure
- Institutional infrastructure

It is intended that evaluation of eGovernment strategy achievements, should be performed by using group of indicators.

The following key factors were identified for accomplishment of eGovernment vision:

- Political willingness
- Leadership and strategic planning
- Human resources
- Financing
- Resistance to the changes
- Participation of citizens and businesses

In the strategy is stressed the necessity of strategic partnership among key players: government, local governments, civil society, businesses, academic/scientific community, and ICT sector.

As a key area of activity, there are stated actions which should be started and which could be organized based on three main area of activity:

- Planning and organizing development of eGovernment
  - Establishing of institutional framework through creation of organizational structure with clear separation of responsibilities for functions of creating policy/strategy, coordination, development, support,

revision/quality control, security/data protection
    - Based on this strategy, the long-term plan of development for the next 5 years should be created
    - Development of eGovernment standards
- Creation of suitable environment for development of eGovernment through:
    - Developing of necessary legislative framework
    - Building of ICT infrastructure
    - Building of security infrastructure – defining and building of PKI
    - Defining and regulating area of e-Payment
    - Improving ICT skills of governmental staff, as well as for experts
    - Promoting of eGovernment
- Development of public services and eGovernment contents
    - Re-engineering and standardization of administrative procedures
    - Development of common services of public data and infrastructural components (services of access to data registries and other available public databases, using defined standards for interoperability)
    - Development of public services government-citizens (G2C) and government-businesses (G2B) (Priority for development has 20 standard basic services)

In the final part of the strategy, there is Action plan for one-year period, which is related to three main areas of activity of eGovernment, i.e. planning and organizing development of eGovernment, creating environment for development of eGovernment, and developing eGovernment services. In that Action plan, there are suggested activities that should be performed, in order to achieve desired goals, and also deadlines for finishing of those activities.

Concrete actions aligned with the Action plan are started in the mid 2006, when Sector plan for investment in eGovernment area as a part of National Investment Plan is accepted. The total amount of investment is 34mil. €, and it is dedicated for funding projects for national computer network, electronic public procurements, improving of local government IS capabilities, and fiscal decentralization.

Currently in Serbia, there is no any special document regarding either strategy for eGovernment, or action plans for eGovernment development.

At this moment (late 2007) it is not fully clear what steps will be taken by the new government in Serbia. Basically, the plan is to form an IT council consisted of IT experts from various ministries. After establishing of current state, the new action plans will be proposed.

4 Legislation Framework

In the Republic of Serbia, regulations related to development of information society can be categorized in two groups: regulations adopted by the year 2001, and regulations adopted in the period 2003.-2005 [6] [5].

Regulations adopted by the year 2001:

1. Law on the Information System of the Republic of Serbia ("Official Journal of Republic of Serbia", Nr. 12/96)
2. Law on Personal Data Protection ("Official Journal of Federal Republic of Yugoslavia", Nr. 24/98, 26/98)
3. Regulation on the Program of Development of IS of the Republic of Serbia in 1997. ("Official Journal of Republic of Serbia", Nr. 3/97)
4. Regulation on the Program of Development and Operability of IS of the Republic of Serbia in 1998. ("Official Journal of Republic of Serbia", Nr. 17/98)
5. Regulation on the Program of Development and Operability of IS of the Republic of Serbia in 2000. ("Official Journal of Republic of Serbia", Nr. 10/00)
6. Regulation on the Program of Development and Operability of IS of the Republic of Serbia in 2001. ("Official Journal of Republic of Serbia", Nr. 58/01)
7. Regulation on Security and Protection of IS of Governmental Institutions ("Official Journal of Republic of Serbia", Nr. 41/90)
8. Manual for Development and Adoption of projects of Information Systems of Governmental Institutions ("Official Journal of Republic of Serbia", Nr. 49/89)

Regulations adopted in the period 2003-2005:

1. Law on Electronic Signature ("Official Journal of Republic of Serbia", Nr. 135/04)
2. Sub-law acts for fulfillment of the Law on Digital Signature ("Official Journal of Republic of Serbia", Nr. 48/05):
    a. Regulation on record keeping of certification bodies
    b. Regulation on register of certification bodies for issuing qualified electronic certificates in the Republic of Serbia
    c. Regulation on technical and technological steps for creating qualified electronic signature and criteria that needs to be fulfilled by the means for creating qualified digital signatures
    d. Regulation on conditions for issuing qualified electronic signatures.
3. Law on Free Access to Public Information ("Official Journal of Republic of Serbia", Nr. 120/04)
4. Law on Business Entities Registration ("Official Journal of Republic of Serbia", Nr. 55/04)

5. Penal Law ("Official Journal of Republic of Serbia", Nr. 85/05)
6. Law on Organization and Jurisdiction of Government Authorities in the Suppression of High Technology Crime ("Official Journal of Republic of Serbia", Nr. 61/05)
7. European Convention on Cyber Crime (2001. Budapest). Serbia signed this convention in April 2005, but it still awaits ratification.
8. Law on Telecommunications ("Official Journal of Republic of Serbia", Nr. 44/03)

In the Republic of Serbia, preparations for development of the Information System of the governmental institutions began in 1989, earlier than in the surrounding countries. By adopting of the Law on Information System of the Republic of Serbia (1996) and the regulations for its fulfillment, legislation framework was created for realization of the significant infrastructural projects on the national level (common computer communication network of the governmental bodies, register of citizens...). However, organizational, technological, human resources and material preconditions for fulfillment of the law were not met on the republic level.

Besides technological, legal regulations adopted in the 90's, today are surpassed from the aspect of actual social requirements for the implementation of the concept of eGovernment and the needs for alignment with the EU and international legislation.

Regulations of the importance for eGovernment were adopted in Serbia in years 2004/2005, later than in the surrounding countries. This especially applies on the Law on Electronic Signature, by which electronic signature is legally made equivalent with the ordinary signature; on the Law on Free Access to Public Information by which paper and digital documents are made equivalent from the aspect of availability, and which offers email communication between citizens and public authorities; on the Law on Business Entity Registration, based on which unique, centralized, public, electronic database on business entities, available over the Internet, was developed in the January 2005. Cyber crime and criminal acts against intellectual property are sanctioned by the penal law, and by special law, specialized governmental bodies were formed for fighting the high-tech crime.

In order to create better legislation framework for more dynamic development of the information society and the eGovernment, it is necessary to adopt new laws from the related areas. This means adoption of:

- Law on eGovernment,
- Law on e-Procurement,
- Law on eCommerce,

as well as other laws necessary for normal process of EU integration and modern world trends. Also, it is necessary to align certain normative acts with the EU directives:

Adoption of amendments and modifications of the Law on Telecommunications from the 2003, by which area of electronic communications should be totally aligned with the so called Second Legislation Framework for Electronic Communications of EU from the year 2002.

- Law on Personal Data Protection, which should guarantee privacy of users of electronic communications, privacy of personal data of employees, privacy of personal data of citizens in the electronic registers of private and public institutions. Also, by this law an independent body for supervision of appliance of this law should be established.
- Ratification of already signed EU conventions such as European Convention on Cyber Crime (Budapest 2001.). This convention was signed by the Republic of Serbia, but it is still not ratified.

Considering the legislation framework, we can say that there is no significant lack in regard to the countries in the region. But, the primary task in the future is finishing of adoption of all necessary sub-acts of Law on Electronic Signature, as a main prerequisite for enabling high-level of electronic public services sophistication. It is announced by the new government that this activity will be finished in 2008.

However, it should be pointed that the bigger problem is the appliance of the existing laws and regulations, as well as finishing of started projects, or putting in the operational state finished projects.

5 Key organization in the development and application of eGovernment in Serbia

In the Republic of Serbia, operations of government administration and tasks in the area of information society development are under jurisdiction of Ministry of Telecommunications and Information Society (jurisdiction taken over from the Ministry of Science and Environment Protection in former government), National Information Technology and Internet Agency, and Office for Common Operations of Government Institutions. Although it is not identical activity as eGovernment, here will be given separation of responsibilities among institutions for information society development, since there is no special structural responsibility regarding eGovernment activities [4] [5].

Responsibilities are divided such that:

1. **Ministry of Telecommunication and Information Society** performs tasks of government which is related to:
   - Creation of policy and strategy for building information society
   - Preparation of laws, other legal acts, standards, and measurement in area of electronic business

- Application of information technology and Internet
- Providing information services

2. **National Information Technology and Internet Agency** (formed in 2003.) performs common and expertise tasks of government related to:
   - Improvement, development, and functioning of information systems of government institutions, local government, and public services
   - Data protection
   - Development and adaptation standards for involving information technologies in government institutions
   - Other tasks regulated by law
   - There are some tendencies that role of this institution will change in the future.

3. **Office for Common Operations of Governmental Institutions,** performs tasks of automatic data processing which evolve:
   - Creation and involvement of projects for automation of administrative and other operational tasks which are performed in Office and other government institutions
   - Design and organization of document and other data bases in charge of government
   - Other information technology tasks related for assurance of functioning and development of information systems

Unfortunately, until now, any wider consultancy mechanisms in order to create forum of stakeholders for discussion of applicability of national strategy for information society, were not developed by the government. There is a noticeable lacking of single institution which should have all necessary permissions for controlling efforts for building information society as well as for eGovernment. Though it is formed with that goal, National Information Technology and Internet Agency did not achieve any significant results, foremost because of lacking funds and conflict of jurisdictions with other ministries.

By looking at the organization of development and functioning of eGovernment, following conclusions can be stated:

Central Government Related Tasks:

- **Policy/Strategy -** In The Republic of Serbia, basic goals and directions of government policy/strategy in the area of information society development and eGovernment development are defined in the document "Strategy for reform of public administration in Serbia", and that strategy is a framework for Strategy for information society development. As described above, currently in Serbia there is no institution on the national level which holds jurisdiction over creating policy/strategy of eGovernment development. National Information Technology and Internet Agency is responsible for improvement, development and assuring functionality of government institutions information systems, local government, and public service offices, which is not the same as eGovernment
- **Coordination -** Currently in Serbia, unlike the other surrounding countries, there is no any multiresor body on national level which is authorized to coordinate and lead the development of the eGovernment in national boundaries. This absence is recognized in the strategy, but there haven't been any steps in this direction yet.
- **Implementation -** For implementation and technical support to the implementation of projects with common importance (network of government institutions, Internet domain of government institutions) as well as giving ICT support to various administration offices is responsibility of Office for Common Operations of Government Institutions, Department for Information Technology, Telecommunication, and Internet. Support to implementation of some projects is provided by IT companies based on outsourcing agreement. Ministry of Science and Environmental Protection is responsible to ensure support of implementation Law on Electronic Signature and it is responsible for establishing security infrastructure of electronic signature.
- **Support** – Responsibility of National Information Technology and Internet Agency and Office for Common Operations of Government Institutions
- **Revision/Assurance -** According to the strategy, State revision institution, as a stand alone and independent institution, liable to National Assembly of Republic Serbia, will perform the revision of government operations, including responsible institutions for ICT.
- **Protection** – Responsibility of Ministry of Telecommunication and Information Society
- **Other -** Currently in Serbia, there are no any authorized institutions-bodies for some specific segments (for example Security and Protection…). Inside of Serbian Chamber of Business and Belgrade Chamber of Business, there is Alliance for Information Activity, which deals with monitoring of activities in ICT area in a business manner. As a part of Standing Conference of Towns and Municipalities there is a council for information technology activities, which is responsible for tracking of state from this field in local governments. Also, there is Information Society of Serbia, as non-profit non-governmental organization. IT council inside of National Investment Plan is authorized

for approving projects related to ICT funded by NIP.

In Serbia, similarly like in other states in region, beside institutions on national level, which have responsibilities for development of common components of eGovernment, other ministries and offices are responsible for development and implementation of their own information systems. For that purpose, the most institutions have special units (sectors, departments, groups) for performing these tasks.

Local Governments Related Tasks:

- **Policy/Strategy –** Responsibility of National Information Technology and Internet Agency
- **Coordination** - Responsibility of National Information Technology and Internet Agency
- **Implementation** - Responsibility of National Information Technology and Internet Agency and Office for Common Operations of Government Institutions
- **Support** - Responsibility of National Information Technology and Internet Agency and Office for Common Operations of Government Institutions
- **Revision/Assurance –** There is no institution or organization assigned for performing these tasks
- **Protection** – Responsibility of Ministry of Telecommunication and Information Society

6 Infrastructure for eGovernment

Following issues are considered as a measure of sophistication of infrastructure for eGovernment services:

- Central Portal of eGovernment services
- Computer network of eGovernment
- Infrastructure for e-Identification
- Infrastructure for e-Procurement
- Infrastructure for knowledge management

Detailed description of state for each issue is given below.

6.1 Central portal of eGovernment services

At the beginning of 2007 the central portal of eGovernment services in Serbia was created, on the address www.eUprava.gov.yu [9]. The portal is created by National Internet and Information Technology Agency. Thereby is achieved one of the goals presented in Action plan for eGovernment in National Strategy for Information Society.

Portal is created with several goals:

- To implement one access point for using existing and future electronic government services
- To join access to twenty basic services by European standards, dedicated to citizens, companies, and public administration

During creation of the central portal for eGovernment, it has used such concepts like user friendly, simple access, with services in front plan. The Portal is made in Serbian and English language, with planned extension on languages of minorities and other world languages. The following standards are respected during implementation:

- Portal is suited to the users needs
- Simplicity and accuracy of presented information
- Service-oriented

Target groups of these portals are citizens, companies, governmental institutions, non-governmental institutions, foreigners who work or are in visit in Serbia.

In future, several activities on eGovernment portal are planned:

- Updating portal content
  - In phase 1: by Agency, with verification from appropriate institutions
  - In phase 2: Directly by institutions responsible for this particular service (after training working on CMS)
- Adding new services

6.2 Computer network of eGovernment

Currently, the national computer network of governmental institutions is implemented just in one part. The project e-Serbia funded by National Investment Plan (NIP) should build unique computer network of governmental institutions in Serbia, and institutions of special importance. The project started during 2006, and it is planned to last for 24 months.

6.3 eIdentification

The Law on Electronic Signature, as the first prerequisite for e-Identification, was adopted in late 2004. After that, during 2005 set of associated sub-acts were adopted, which regulated area of work for certification bodies for issuing digital signatures. But not all necessary sub-acts are adopted yet. Because of problem in enforcement of these acts, and also due to absence of other laws (Law on Personal Data Protection), there is no yet national certification body in Serbia.

However, from 16.11.2004 Post Serbia Certification Authority is issuing (selling) digital certificates to interested users outside of Post Company, no matter if they are individuals or legal entities. At this moment, Post Company of Serbia, actually Post Serbia Certification Authority is the first and the only public certificate body in the Republic of Serbia [7].

Inside Ministry of Interior, the project of Electronic Personal Cards (and also e-Passport project), run in cooperation with the Agency for Printing Banknotes, is in the final phase.

By Law on Personal Cards (2006), citizens are allowed to choose between standard and electronic personal cards. The project eCards has started in 2001 and is delayed mostly because of technical problems, as well as for dissatisfaction of other requirements (infrastructural, legal, absence of certificate bodies...) for full application of such project.

The type of data that could be stored on eCard is strictly regulated by the law. On the other hand, before beginning of using eCards, it is necessary to adopt the new Law on Data Protection, which should be adjusted with EU recommendations.

Together with the eCards project, the ePassport project is running.

But, as noted in section of Legislative Framework, the main barrier for development of efficient eIdentification infrastructure is lack of all necessary sub-acts of Law on Electronic Signature.

6.4 eProcurement

The main barrier to the implementation of electronic public procurements is of legal nature, i.e. it is necessary to adapt (or to supplement) Law on Public Procurements (adopted in 2002. with changes in 2004) with appropriate normative and organizational solutions for full online utilization.

Inside of National Investment Plan the e-Procurements project was approved, and should be run during 2007. It is the common project of Ministry of Finance and National Internet and Information Technology Agency, with participation of Public Procurement Office, and Ministry of Telecommunication and Information Society.

The goals that software system e-Procurements should satisfy are (similarly like in other countries where this system is already implemented):

- Increasing of effectiveness, efficiency and transparency in process of public procurements
- Decreasing corruption in governmental institutions
- Shorter time with lower expenses of procurement process better concurrency, with lower prices and faster economical development procurement process could be monitored from any physical place, in any time

In conduction of public procurement through the internet portal, terms of adapted Law on Public Procurements will be respected and will enable that every operation that was made from the computer of responsible person accomplish safely. It is expected that project e-Procurements facilitate faster assurance faithful certificate body of governmental institutions, local governments, and public administration for their need of electronic business.

Web address of Public Procurement Office is http://www.ujn.sr.gov.yu.

6.5 Infrastructure for Knowledge Management

At this moment, there is no infrastructure for Knowledge Management in Serbia on governmental level.

7 State of basic eGovernment services in Serbia

7.1 Introduction

In the following analysis, there are used indicators defined by EU Commission (benchmarks) for a measurement of sophistication of electronic public services in Serbia. Here will be covered just the area of eGovernment, actually the basic 20 services of eGovernment: 12 services for citizens, and 8 services for companies [2].

Public services for citizens involve:

1. *Income Tax Declaration*
2. *Job Searches by labour Offices*
3. *Social Security Contributions*
4. *Personal Documents (passports and driving license)*
5. *Car Registration*
6. *Application for Building Permission*
7. *Declaration to the Police*
8. *Public Libraries*
9. *Certificates (birth, marriage) Request and Delivery*
10. *Enrolment in Higher Education*
11. *Announcement of moving (change of address)*
12. *Health-related services (e.g. appointments for hospitals)*

Public services for companies involve:

1. *Social contributions for employees*
2. *Corporation Tax: declaration, notification*
3. *VAT: declaration, notification*
4. *Registration of a new company*
5. *Submission of data to statistical offices*
6. *Customs declaration*
7. *Environment-related permits*
8. *Public procurement*

As a measuring indicator of public services, the same four-stage framework defined in [1] is used:

- Stage 1- Information: The information necessary to start the procedure to obtain this public service is available on-line.
- Stage 2- One-way Interaction: The publicly accessible website offers the possibility to obtain in a non-electronic way (by downloading forms) the paper form to start the procedure to obtain this service. An electronic form to order a

non-electronic form is also considered as stage 2.

- Stage 3- Two-way Interaction: The publicly accessible website offers the possibility of an electronic intake with an official electronic form to start the procedure to obtain this service. This implies that there must be a form of authentication of the person (physical or juridical) requesting the services in order to reach stage 3.
- Stage 4- Full electronic case handling: The publicly accessible website offers the possibility to completely treat the public service via the website, including decision and delivery. No other formal procedure is necessary for the applicant via "paperwork".

Stage 0 has been used for ranking either total absence of any publicly accessible website managed by the service provider, or the case when the public service provider has a publicly accessible website, but this one does not offer any relevant information, interaction, two-way interaction or transaction possibilities at all concerning the analysed service.

7.2 Public Services for citizens

Below is given description of each 12 services for citizens, with ranked stage of sophistication [5].

7.2.1 Income taxes

Responsibility: **Tax Administration, Ministry of Finance**
Web site: **http://www.poreskauprava.sr.gov.yu**
Sophistication: **50%**
Description of service: Regulations, guidelines and electronic forms are available for downloading

7.2.2 Job search

Responsibility: **National Employment Service**
Web site: **http://www.rztr.co.yu**
Sophistication: **75%**
Description of service: There are job searching facilities, submitting application in the database of employees, submitting application in the database of employers

7.2.3 Social security benefits

a. Unemployment Benefits
Responsibility: **National Employment Service**
Web site: **http://www.rztr.co.yu, http://www.euprava.gov.yu/servisi/2_4_10/3_4_10_1/**
Sophistication: **25%**
Description of service: There is online available information about rights on allowance for unemployed

b. Family allowances
Responsibility: **Ministry for Work and Social Policy, Local Governments**
Web site: **Sites of municipalities, http://www.minrzs.sr.gov.yu, http://www.euprava.gov.yu/servisi/2_4_10/3_4_10_3/**
Sophistication: **25%**
Description of service: Local governments are responsible for this service. On their sites are located information about necessary documents and procedures.

c. Medical Costs
Responsibility: **Ministry of Health, National Agency for Health Insurance**
Web site: **http://www.rzzo.sr.gov.yu**, **http://www.euprava.gov.yu/servisi/2_4_10/3_4_10_2/**
Sophistication: **25%**
Description of service: Online information about reimbursement of medical costs

d. Student grants
Responsibility: **Ministry for Education and Sport**
Web site: **http://www.minrzs.sr.gov.yu**
Sophistication: **25%**
Description of service: Information about applications for student grants

7.2.4 Personal documents

a. Passports
Responsibility: **Ministry of Interior**
Web site: **http://www.mup.sr.gov.yu**
Sophistication: **33%**
Description of service: Information about necessary documents, charges

b. Driver's license
Responsibility: **Ministry of Interior**
Web site: **http://www.mup.sr.gov.yu**
Sophistication: **33%**
Description of service: Information about necessary documents, charges

7.2.5 Car registration

Responsibility: **Ministry of Interior**
Web site: **http://www.mup.sr.gov.yu**
Sophistication: **25%**
Description of service: Information about necessary documents, charges

7.2.6 Building permission

Responsibility: **Ministry of Capital Investments, Local government offices responsible for urbanism, building of premises, and residential relations**
Web site: **http://www.mki.sr.gov.yu**, **http://www.euprava.gov.yu/servisi/2_4_7**
Sophistication: **25%**
Description of service: Information about regulations, procedures, and necessary documents

7.2.7 Declaration to the police

Responsibility: **Ministry of Interior**

Web site: **http://www.mup.sr.gov.yu/**, **http://www.euprava.gov.yu/servisi/2_4_9**
Sophistication: **33%**
Description of service: There is online information about starting procedure of declaration to the police.

7.2.8 Public Libraries

Responsibility: **Ministry of Culture, Public Libraries**
Web site: **http://vbs.nbs.bg.ac.yu/cobiss/**, **http://www.biblioteke.org.yu**
Sophistication: **75%**
Description of service: In the virtual library network of Serbia, it is possible to get information about manuals, catalogs, reading rooms, book borrowing, citation, working rules, charges, making a reservation for a book is possible only in a few libraries

7.2.9 Certificates

Responsibility: **Local Governments, Ministry for Public Administration and Local Governments**
Web site: **Sites of municipalities, http://www.mpalsg.sr.gov.yu**, **http://www.euprava.gov.yu/servisi/2_4_1**
Sophistication: **66%**
Description of service: There is possibility to download form for starting non-electronic procedure. In some municipalities, there is possibility for online submitting requests for getting birth, marriage, or voting right certificate. Implementation of the service varies on concrete municipality.

7.2.10 Enrolment in higher education

Responsibility: **Ministry for Education and Sport (Department for High and Higher Education), Universities**
Web site: **Sites of Universities, http://www.mps.sr.gov.yu**
Sophistication: **25%**
Description of service: On the sites of some faculties, one can get information about enrolment conditions, educational program

7.2.11 Announcement of moving

Responsibility: **Ministry of Interior**
Web site:**http://www.mup.sr.gov.yu, http://www.euprava.gov.yu/servisi/2_4_2**
Sophistication: **33%**
Description of service: Information of necessary documents, charges

7.2.12 Health related services

Responsibility: **Ministry of Health, Hospitals, Clinics, Health consulting-rooms**
Web site:**http://www.zdravlje.sr.gov.yu, http://www.euprava.gov.yu/servisi/2_4_5**
Sophistication: **25%**
Description of service: On the sites of some health institutions (especially private ones), there is a possibility for making appointments

7.2.13 Conclusion

The following table illustrates online sophistication for citizen services in Serbia comparing with EU average sophistication. It is clear that Serbia is far away from EU average. The full enforcement of Law on Electronic Signature could significantly improve this result without bigger technological change in short period.

Serbia 2007	EU(18) 2006	EU(10) 2006	EU(28)
40.83%	71.00%	62.00%	68.00%

Table 1 - Citizen Online Sophistication: Serbia 2007 and EU 2006

7.3 Public services for businesses

7.3.1 Social contributions

Responsibility**: National Fund for Pension and Invalidity Insurance of Employees, Tax Administration of the Ministry of Finance**
Web site:**http://www.euprava.gov.yu/servisi/2_5_4 http://www.rzzo.sr.gov.yu /, http://www.poreskauprava.sr.gov.yu/**
Sophistication: **50%**
Description of service: Online information about necessary procedures, forms could be downloaded

7.3.2 Corporate tax

Responsibility**: Tax Administration, Ministry of Finance**
Web site:**http://www.poreskauprava.sr.gov.yu**, **http://www.euprava.gov.yu/servisi/2_5_3**
Sophistication: **50%**
Description of service: Electronic forms could be downloaded from the site. There is a special treatment for big tax tributary

7.3.3 VAT

Responsibility: **Tax Administration, Ministry of Finance**
Web site:**http://www.poreskauprava.sr.gov.yu/ http://www.euprava.gov.yu/servisi/2_5_5**
Sophistication: **50%**
Description of service: Electronic forms are available for download. Various information about VAT and procedures for charging VAT are available.

7.3.4 Company registration

Responsibility: **Serbian Business Register Agency**
Web site: **www.apr.sr.gov.yu**
Sophistication: **50%**
Description of service: Download of electronic forms necessary for company registration, and forms for modification of existing information in the registry. There is an opportunity for searching of business registry and also other registries in responsibility of this agency.

7.3.5 Statistical data

Responsibility:**Statistical Office of the Republic of Serbia**
Web site:**http://www.statserb.sr.gov.yu, http://www.euprava.gov.yu/servisi/2_5_2**
Sophistication: **66%**
Description of service: One can download various forms from the area of living environment, population, architecture

7.3.6 Customs declarations

Responsibility: **Custom Administration**
Web site: **http://www.fcs.yu**
Sophistication: **75%**
Description of service: Ability for electronically submitting declarations (after downloading client application), information about regulations, manuals, and electronic forms are available online, interactive completing of JCI

7.3.7 Environment-related permits

Responsibility: **Ministry of Environmental Protection**
Web site: **http://www.ekoserb.sr.gov.yu**
Sophistication: **25%**
Description of service: On the site there are available regulations, manuals, and reports which could be downloaded. There are no forms for download.

7.3.8 Public procurement

Responsibility: **Public Procurement Office**
Web site: **http://www.ujn.sr.gov.yu**
Sophistication: **75%**
Description of service: Advertising of public procurement, defining criteria for the best possible offer, gathering offers, making decisions about requests with additional explanations, opening and ranking of the most acceptable offers; negotiating, and monitoring the process through to completion, and statistical reports. Electronic identification of participants, uploading and downloading of electronic documentation and reports, submitting and answering the questions, offer ranking, publishing final results etc.

7.3.9 Conclusion

Overall online sophistication for business services in Serbia and in EU countries is given in table below. It is clear that all EU countries gave a stress on this kind of services, but Serbia has to improve a lot.

Serbia 2007	EU(18) 2006	EU(10) 2006	EU(28) 2006
55.13%	88.00%	81.00%	85.00%

Table 2 - Companies Online Sophistication: Serbia 2007 and EU 2006

7.4 Cluster Online sophistication

7.4.1 Income-Generating Cluster

This cluster involves the following services: taxes, social contributions, VAT, and customs.

In this cluster, that addressed services for businesses we see that average level for EU countries is very high, but in Serbia is just above 50%. Improving sophistication of these services will definitely enable better environment for business in Serbia.

Serbia 2007	EU 2006
55.00%	94.00%

Table 3 – Income-Generating Cluster Sophistication: Serbia 2007 and EU 2006

7.4.2 Registration Cluster

This cluster involves the following services: car registration, company, birth & marriage, moving, and statistical data.

Serbia 2007	EU 2006
48.00%	72.00%

Table 4 – Registration Cluster Sophistication: Serbia 2007 and EU 2006

7.4.3 Returns Cluster

This cluster involves the following services: health, libraries, procurement, policing, job search, and benefits.

Serbia 2007	EU 2006
51.33%	71.00%

Table 5 – Returns Cluster Sophistication: Serbia 2007 and EU 2006

7.4.4 Permits and Licenses Cluster

This cluster involves the following services: building, passport, education, and environment.

This result show that in Serbia is very low level of sophistication for online services regarding getting some permits or licenses, and these services should be improved.

Serbia 2007	EU 2006
27.00%	61.00%

Table 6 – Permits and Licenses Cluster Sophistication: Serbia 2007 and EU 2006

8 Services important for transactional cross border application domain

8.1 Introduction

It is noticed that in Western Balkan Countries there is a low level of awareness of interoperability issues and cooperation potentials. In order to enable public administration in Western Balkan Countries (WBC) to reach higher productivity and equity by establishing of an eGovernment Interoperability Framework with focus on transactional cross border services, here will be given description of current state in governmental institution responsible for such services, important for transactional cross border application domain.

The list of these important cross-border services include eCustoms - New Computerized Transit System (NCTS), eFinance - VAT Information Exchange System (VIES), eJustice cross border cases - Automation of Court Procedures (ACP), European Companies Register (ECR) and European Land Information System (EULIS), eAdministration - Electronic Filing System (implementation of paperless government) and eTrade Facilitation for European Waste transport (EUDIN).

Please, keep in mind that this description is based on publicly available information about these services.

8.2 eJustice

The National Judicial Reform Strategy in Serbia has been adopted in April 2006 [10]. This Strategy sets forth the challenges facing Serbia's judiciary within the framework of four key principles and corresponding goals. An effective justice system is based on four key principles: independence, transparency, accountability, and efficiency, which provide the framework for the design, development and organization of all judicial institutions. A separate Implementation Plan outlines the specific steps needed to achieve these goals.

The Development and Implementation of the Judicial Information System of Serbia (JISS) [11] is under jurisdiction of the Ministry of Justice according to law. JISS includes 7 Information Systems (IS): IS of the Ministry of Justice as a central point of JISS, IS of courts of general jurisdiction, IS of Commercial courts, IS of Misdemeanor courts, IS of the Administration court, IS of Public Prosecution Offices, and IS of Prisons.

Recently, several projects have been done, or are still in implementation, related to court modernization. Some of those projects are funded by EU, some with other donors, and some are funded locally [12] [13] [14]:

- Several pilot projects are implemented recently: involving of CMS in the First Public Prosecution Office in Belgrade, Special Court in Belgrade is a project funded by USAID and managed by NCSC, Municipal court in Zajecar is the project funded and managed by the German IRZ foundation, District Public Prosecution Office in Belgrade, example of involving of CMS in the biggest PO in the Serbian Judiciary.
- In cooperation with European Agency for Reconstruction (EAR), the second phase of the project "Supporting Judicial Reform in Serbia" was finalized. The project encompasses 5 judicial institutions including district and municipal courts, district and several municipal public prosecution offices. Work on improving information infrastructure has been completed. Also, with the help of EAR, SENA business software has been prepared for usage in courts of general jurisdiction. The software is based on Microsoft platform, and it offers possibility for electronic management of court records, tracking status of cases, automatic collection and processing of statistical data, as well as preparation of data for annual reports on judges' efficiency. It was expected that SENA should be installed in some courts by the beginning of March 2007.
- Commercial Court Administration Strengthening Activity (CCASA) is a three-year, $12.8 million project began in March 2004, funded by the United States Agency for International Development (USAID). Up to date, all commercial courts (The Higher Commercial Court and 18 Commercial courts) are connected in a computer network (VPN). CAS - Court Automation System in three modules was introduced. Web-based statistics gathering software known as Case Data Collection Instrument (CDCI) is implemented in the entire Serbian Commercial Court system. In year 2006, all cases (over 117.000) were entered using CDCI; Currently, statistical reporting is available for more than 183.229 cases
- Information on commercial courts, including court statistics, court schedules and data on bankruptcy cases available to the public on the web portal: www.trgovinski.sud.srbija.yu;
- Support for the project "Court administration modernization based on IT CMS – CASE MANAGEMENT SYSTEM" was finalized in cooperation with German IRZ foundation (http://www.irz.de). The result was start of a pilot project – Registry Office in the form of service units for the operation of the business software in Belgrade District Court and 8 municipal courts.
- "Database of legal regulations and court practice" project was finished in cooperation and with the donation of European Agency for Reconstruction – EAR. Since July 2006, electronic database is fully functional and can be used over the network of governmental institutions.
- Ministry of Justice in cooperation and with the donation from the Kingdom of Norway is working on a project "Justice Network - forensic

scientists and interpreters". The goal of this project is to, in shortest time possible, make publicly available basic information on courts in the Republic of Serbia.

- The project "Supporting Crime Services in Serbia and Montenegro" was finalized in February 2007, and it encompasses all public prosecution offices in the Belgrade area, which are connected in a common computer network with the center in the Palace of Justice in Belgrade.
- Software for submitting complaints of the parties was installed in the Section for Court Supervision in the Ministry of Justice. The software is used for submitting all complaints towards Ministry of Justice; it has searching and analysis capabilities. Till now, total of 11 647 cases were processed using this application.
- Development of the Information System of the Ministry of Justice, is funded and managed by the Microsoft Office in Belgrade

Ongoing modernization of the information technologies (planned for period 2007-2011 according to the implementation plan of the national strategy) will enable monitoring of courts' and judges' efficiency, monitoring their workload, access to electronic database of legal regulations and court practice, efficient way of collecting all statistical data, and interactivity with the public.

Here are some plans or ideas for the next period:

- The project named "ePravda" (eJustice) is submitted to the government for funding. This project is planned to last for three years and to solve many lacks that are present today.
- UNDP support the Misdemeanor courts and support for the Judicial Training Centre. Inside the working group for the realization of the donating project UNDP „Improving Penalty System and Misdemeanor Courts in the Republic of Serbia", expert team was formed with the assignment to make the evaluation of existing software solutions for monitoring the work of misdemeanor institutions.
- Complete analysis of the current state of IT infrastructure, and preparations for the plan for the realization of the project "IT modernization of the court institutions" from the National Investment Plan were performed. Realization of this project is planned for the year 2007.
- It is expected improvement of IS of Prisons and The Administration court
- Also, there is a plan for sharing of business and technical data in the Judicial Information System of Serbia.

Generally, in the Serbian Justice system there are some particular solutions that are very modern, as a result of international donations. However, overall impression is that system is not fully integrated, not standardized, and therefore not ready for delivering various (especially complex) services. Also, communication with other governmental institutions is very poor. Practically, Justice Information System stands as an isolated system. There is a lot of room for interoperability here. It seems that the proposed project ePravda (eJustice) will tend to solve many of these problems.

8.3 eFinance

Ministry of Finance is responsible for some services important for cross-border applications. Here will be given a description about related institutions inside of Ministry of Finance.

8.3.1 eTreasury

Currently, there are a lot of activities related to information system which is in function of controlling public finance. The name of the main project is Integrated Information System of the Treasury (FMIS) [15]. The project is divided in four phases and is planned to finish by the end of 2008.

The key goals of this project are:

- Implementation of integrated IT solution in charge of Treasury, relying on the most modern technologies, in concordance with relevant international standards, due to optimal and effective controlling of Consolidated Account of Treasury. Also, it should be continued development and modernization of accounting, executing and control of budget, as well as establishing of sophisticated function of financial planning and anticipation.
- Constitution of electronic service access, in online regime, to the all Direct Budget Users (DBU), and Indirect Budget Users (IBU) for completing process of budget execution in which it will be no more communication with papers, whereby it will significantly be improved control and transparency (visibility) of the whole process.
- Implementation of the system for controlling debt and liquidity in order to make feasibility of Serbia for effective and optimal debt management through integrated IT solution, according to the best world practices and standards.

Implementation of these goals requires creation of new business processes for DBUs and IBUs, replacement of temporary treasury solution, creation of the new tools for queries and reports (execution of the budget, accounting), and training of DBUs human resources and change management.

8.3.2 eTaxes

In the area of eTaxes there is a final phase of the big project funded by National Investment Plan. It is the project Fiscal Decentralization in Serbia (FIDES). The key players in the project are Ministry of Finance (Tax Administration) and National Information Technology and Internet Agency. However, Ministry of Public Administration and Local Government also participate in the project, as well as Ministry of Science and Environment Protection.

Strategic goals of this project are related to:

- Efficient work of local governments
- Exchange data with other information systems in order to achieve unique relevance of citizen and assets data
- Continuity and higher efficiency in the business of validation and charging of profit

Operational goal of the project is creation of the system for support to the fiscal decentralization of tax system on the level of local governments with direct exchange of data with central tax administration and Ministry of Finance.

Future local tax administration will determine, control and charge tax on assets of individuals and legal entities, tax on transfer of absolute rights, tax on inheritance and gifts, tax on income of individuals' real estate, local communal tax, imbursement for using and arrangement of plot of ground, imbursement for protection and improvement of living environment, residence tax, voluntary tax, and other public income regulated by law.

The final goal of this project is financial decentralization and enabling local governments to more independently decide about its local economical development.

However, there are several continual goals that are planned to be achieved during implementation of this project:

- Development of local governments with optimal and consistent tools for enforcement of own local competency in order to do faster implementation of relevant EU recommendations and standards in domain of local governments
- Efficient administration and charging of local tax income wherewith will improve local budgets and decrease demand on central government budget
- Building of national PKI in governmental institutions and local governments as a basic requirement for secure finalization of this project
- Building of national system for relations with clients – Calling / SMS / E-mail Center. For each of the previously quoted functions, it will be introduced two-way communication with clients via phone, SMS, e-mail and optionally by fax.

It is expected that this project will be finished during 2007. Current sophistication of online tax services is 50%(both for citizens and for companies) and it is expected that results of this project will improve quality of service in this area.

8.3.3 eCustoms

Serbian Customs Information System is established in 1974, but in this form works since 1995. SCA IS is a complex and integrated system and supports all basic customs procedures [16].

Structure of the Serbian Custom Administration is as follows: approximately 2500 Serbian Customs officers are located at the Headquarters in Belgrade, 13 customs houses and over 140 customs posts/units/points use the IS in line with their legal powers, upon which over 100.000 computer transactions per day are performed on local servers in Customs Offices and on the central server in the Customs Headquarters.

Hardware platform for Serbian Custom Administration Information System consists of central server (IBM Z800), and local servers (IBM AS/400 – i-Series) connected in a star network. The system is functional 24hours/day, 365days/year.

Custom officers in Customs Houses and their organizational units process all the basic customs documentation on local servers (Single Administrative Document (SAD), TIR Carnet, ATA Carnet, Summary Declaration). After finishing the processing, data is sent to the central server in different intervals (each 2 to 1 minutes a day, depending on the nature of data). Local servers are also used as application servers, because in case of loosing connection with central server they can continue to work independently, because all the needed consultant data and programs are on local servers.

SCA IS consists of a few subsystems which are divided into approx. 100 modules which incorporate a few hundred options that use around 2500 applications:

- General subsystem, which involves user management, reference data management, network configuration, data replication service management, backup management.
- Regulations subsystem: legislation data base, tariff management, customs valuation.
- Transit subsystem: declaration management, confirmation that goods reached destination, processing and management (comparing data, processing unfinished transit operations, etc.), examination report management, road tax administration.
- Clearance subsystem: declaration management, examining and releasing of goods, sampling and laboratory analyses, payment management, calculation of interest, bank guarantee management, and automatic management of customs officers selected for goods inspection.
- Administrative and offence procedures including requests and permissions for customs procedures with economic impact

- Human resource management: organization and systematization, staff evidence etc.
- Statistics subsystem: passengers' data, vehicles data, offences, etc.
- Customs Passenger Declaration: processing the declaration, calculating and charging customs duties for goods that passengers carry with them
- Risk Analysis subsystem: risk analysis system management
- Direct Traders Input – DTI subsystem

Future development of SCA IS is defined in the SCA IS Development Strategy document. SCA IS Development Strategy consists of several parts related to:

- IT Infrastructure
- Implementation of the new functionality
- Upgrade of the existing SCA IS to the new technologies based on open platforms
- Implementation of EU system
- Implementation of support system
- Single Electronic Window, as a crucial goal of SCA IS functionality

One section of SCA IS Development Strategy addresses following issues related to strategic business requests:

- Single Electronic Window
- Transit
- Clearance Procedures
- Enforcement and Compliance Functions
- Road Taxes
- Mission Support Procedures.

Recommendation no. 33 UN/CEFACT, describes Single Electronic Window (SEW) as a facilitation measure for parties involved in trade and transport which enables submission of standardized information and documentation, at a single place, in order to fulfill all requested conditions related to import, export and transit. Also, this concept enables efficient cooperation between different governmental agencies.

Currently, there is a process of SEW implementation that consists of several activities:

- Ensuring close cooperation between governmental agencies, changing the technology of
- Carrying out customs procedures, standardization of data and documents used in the SEW, and IT support to the new form of cross government agencies cooperation and cooperation with the trade community.

The next section in the strategy related to business strategic requests is the Transit section. SCA IS will support all changes in transit procedures during the process of adjusting to the transit procedures that are being used in the customs administrations of the developed countries. System will also support the EU NCTS (New Computerized Transit System) which should be implemented just before Serbia enters EU.

The strategy in section related to Clearance Procedures says that SCS IS will support all changes of clearance procedures in process of adjusting to the procedures that are used in customs administrations of developed countries. System will also incorporate TARIC (integrated EU tariff), which will be the referent database for implementing all customs procedures and automation in calculating taxes that are being charged in customs procedures.

Regarding Enforcement and Compliance Functions, IT support will be provided for intelligence analytics, offence case tracking, disciplinary measures and appeals, risk analysis and audit control. Up to date, it is realized support for offence case tracking, audit control (CAFAO donation), and intelligence analytics (CAFAO donation – insufficient). There is no support yet for intelligence analytics (System for business informing), risk analysis (long term work), and disciplinary measures and appeals.

According to the strategy, SCA IS will support all changes in calculation and collection of Road Taxes. Currently, IT connectivity is agreed between TERTERM system that is used by Public Company Roads of Serbia (and it was developed by Mihailo Pupin Institute) and SCA IS on all border crossing points.

At the end, the strategy addressed issues of Mission Support Procedures. There is a plan for providing support for human recourses management, document data flow management, management and governance support system, data warehousing, analytics and reporting, the SCA Intranet and the SCA Web site.

8.3.4 eFinance conclusion

We may notice that in the public finance there are a lot of activities in the IT area. But, sophistication of their online services is very low comparing with EU average results.

Through participation in various EU projects, Ministry of Finance created its own strategy for IT development, good practice from EU countries is accepted, also an interoperability issues are known and in some areas are followed in order to be connected with relevant EU organizations. But, until now, there is no significant work done in communication with other governmental institutions in Serbia in order to offer high sophistication level of online services.

8.4 Companies Registers

For the purpose of reform of public administration, one of important reform laws was adopted in May 2004, Law on registration of business entities. Based on this law, Serbian Business Registers Agency (SBRA) was formed, as the unique institution responsible for maintaining business registries, and its work is regulated in detail by special Law on Agency for business registries [17].

Start of agency is dating back in 2004 when basic team was formed. Already in September 2004 the project Reform of Business Subjects (project REPS) was started

in collaboration with World Bank. But official start of agency is 4.1.2005 when maintenance of Registry of Business Entities and Registry of Financial Leasing has started.

In August 2005 maintenance of Registry of Piedge Rights has started, and from January 2006 Registry of Businesses, and Submitting Financial Reports of Business Entities (from January 1st 2006). It is expected that during next years this list will be extended with new registries, whereby in shortest time it is expected to activate the Registries of Foreign Investments, and Registry of Factoring.

Besides Serbian government providing of financial facilities especially in preparation phase, was highly supported by: Swedish Govermnent, USAID, and Microsoft. The procurrement procedure was accorrding to the World Banks International Competitive Bidding (ICB) rules. It was „Turn-Key ICT Project", which involved delivery of complete solution for SBRA, including hardware, software and all necessary services. In preparation of tender documentation, important role was played by IT consultants from Ireland.

During project development, a good model from EU was selected, whose experiences were used. That was an Irish model, and a very successful example of knowledge transfer from Companies Registration Office (CRO) Ireland, Dublin.

Modern information system in the SBRA, throughout its iternal part, user/partner connections, and Internet site, enables:

- Economy and data availability through the unique centralized database for every of needed registries
- Simplification and speeding up of registration process
- Decreasing of registration and business costs
- Harmonization with EU standards and directives
- Education of users and improvement of inter-communication
- Electronic data exchange with public and private sector

Before reform, the registration system was indented for: registration of company was performed in one of 17 commercial courts, and registration of businesses in one of 161 municipalities. Automation of the process in the courts was on low level, mostly on the level of very primitive information system, without inter-connection. Automation of the process in the municipalities varied, ranging from solid application down to totally manual processes, without any standardization.

During development of Registry of Business Entities in SBRA (the project REPS), the modern concepts were used:

- Electronic register
  - Centralized, unique database of business entities
  - Usage of the most modern technologies (multi-tier Web applications, RDBMS – MS SQL server, LAN/WAN (Intranet))
  - Full data availability (24/7) – Using of Internet for searching purposes
  - Implementation of data exchange and business collaboration with other institutions
  - Example of eGovernmnent services(G2C, G2B, G2G), done by all standards
- Ability for registration through Internet
  - Ability of downloading of all forms via Internet, and filling of Web forms
  - As soon as it is possible, it will be enabled an offer of full electronic registration with use of electronic signature
- Respect of international standards
  - Concept and legislation framework are aligned with "EU Best Practices"
  - The modern standards of project management and software engineering methodologies (Agile Methods) were used
  - Data formats and communication formats are such that can enable international collaboration

Essential functions of software for Registration of Business Entities are:

- Electronic submission of requests
- Receiving electronic slips
- Scanned documents will be available in electronic way and opened for searching
- Possibility for making payments electronically (EUR, Dinars)
- There are various opportunities for connection based on industrial standards (Web Services based on XML)

Implemented solutions are enabling data exchange with other institutions and offices. The final goal is realization of "one stop shop" services for citizens, throughout several phases.

Presently, there is collaboration of SBRA with the following institutions:

- National Agency for Statistics (from 2004).
  - Exchange of data set related to companies and businesses / Statistical analysis
- National Bank of Serbia
  - Exchange of data set related to companies and businesses / Data on banking accounts
- National fund for pension and invalid insurance
  - Forms of businesses application (electronic), Confirmation of application (paper)
- Ministry of Finance, Tax Administration
  - Data set of companies and businesses / Data of Tax IDs
- Leasing companies
  - Data of leasing contracts / Reports from Registries of financial leasing

Collaboration or extension of collaboration is planned with local governments (municipalities), National agency for health insurance, National agency for employment (sending data of businesses), Custom Administration, and with other agencies and ministries.

Collaboration with private sector is planned also: electronic delivery of data for legal entities and citizens, collaboration with banks (intensive collaboration already exists in some domain), collaboration with Chamber of Business, Alliance of Banks (negotiations are active), commercial registries, enterprise clients (Telecom Company, Post Company, etc.).

It is planned to achieve full electronic registration – for partners and end-users. At the moment the full realization depends on enforcement of Law on Electronic Signature.

The Serbian Business Registers Agency has recently joined the European Business Register (EBR) group. The work for the technical integration to the network will start soon. It is hoped that company information from Serbia will be available by the end of 2007.

In future, there are lots of planned activities that should be done in the SBRA.

Here are given planned activities for the year 2007:

- Finalization of the project "One Stop Shop for registration in SBRA"
- Participation in the BRITE project
- Continuity of quality improvement, standardization of data in the registries, initiative for changing necessary regulations
- Finishing of businesses databases
- Further opening to the public, through complete offer of deliverable electronic slips (complex reports from database)
- Participation on European Commerce Registers Forum (ECRF) in Riga, Latvia; Active international and especially regional collaboration, in the field of registration and more
- Realization of electronic requests (depending on availability of qualified electronic signatures)

Our impression is that SBRA is currently the leading institution in the country by achieved technological progress and potentials for further development.

8.5 eAdministration

Present situation at Public Administration and Local Government in Serbia is such that implemented information solutions and services are varying from municipality to municipality in sense of functionality, contemporary of used technologies and equipment.

In the most number of municipalities, there are information systems that partly implement most important business operations and which are relatively modern and non-integrated.

Used equipment is only partly modern. There are no examples of interoperability or any other attempts of process standardization, automation, etc.

Variety in efficiencies of implemented solutions used by organs of local governments, directly reflect on variety in development of basic eGovernment services such as Application for Building Permission and Certificates (birth, marriage) Request and Delivery that belong to domain of local government responsibilities.

In order to organize this domain, Local eGovernment project has been prepared and it will be financed from resources of National Investment Plan. Preliminary plan was to start project development in the mid 2007, but there is a possibility of prolonging this term.

The basic idea of this project was improvement of efficiency and standardization of providing eGovernment services across entire Serbia, with minimum expense and elimination of multiple data sources, all strictly according to relevant laws.

Also, very important idea and purpose of this project is establishing technical environment with capability of providing reliable and safe communication between local Governments and using systems for temporary transfer of responsibility, based on mutual authentication, using national PKI infrastructure.

The project was proposed on behalf of National Internet and Information Technology Agency, and other project participants are Ministry for Public Administration and Local Governments, and Ministry for Telecommunication and Information Society.

Main functionalities of the project will be: issuing birth, marriage, death and citizenship certificates, Registry Office (Business Process Management / Document & Workflow Management), parliament affairs with support for councilor public relations, support for assets(Property)-Legal's affairs, inspection affairs, urban planning and construction (with legalization), protection for persons with special needs (children, mothers-pregnant women, invalids, family allowances, etc.), human resources management, system for managing local government regulative, Internet/Intranet portal as system for communication with clients, providing preconditions for implementation of election register, payment portal.

The most important goals of the project are: focusing on treating users of Government Administration as clients, improvement of efficiency and transparency, standardization and minimization of provided service expenses. These goals are to be implemented throughout providing services based on concept „to all citizens in all municipalities"

One of the continual goals of the project is development of national system for relations with clients (citizens and enterprises) - Call/SMS/E-mail Center. For each of the specified functions this implementation brings improvements through two-way communication with clients using phone, SMS, E-mail and optionally fax. Special phone numbers will be available for providing such services.

9 Conclusion

If we have to define what the main barrier in Serbia is for providing high level online services, we would like to say that it is interoperability on national level, besides all legislation problems. Actually, there is a lack of all aspects of interoperability: organizational, semantic, and technological. It is strange that for country that it is not in EU, some institutions are better connected with relevant EU institutions, than with local governmental institutions.

We hope that the new team in Ministry of Telecommunication and Information Society will significantly improve general state of eGovernment services in Serbia.

References

[1] Capgemini (2005) Online availability of public services: How is Europe progressing, European Commission Directorate General for Information Society and Media

[2] Capgemini (2006) Online availability of public services: How is Europe progressing, Web Based Survey on Electronic Public Services, *Report of the 6th Measurement*, *i2010* Information Space Innovation & Investment in R&D Inclusion

[3] Ministry of Public Administration and Local Self-Governments of the Republic of Serbia (2004), Strategy for reform of public administration in Serbia

[4] National Agency for Information Technology and Internet (2006), Strategy for development of information society in Serbia

[5] National Agency for Information Technology and Internet (2006), Long-term plan for e-Government development in Serbia, Draft 2

[6] National Information Technology and Internet Agency (2006), Analysis of legislative framework for e-Government in Serbia

[7] D. Spasic (2006), Digital certificate as digital identity of Internet users, Serbian Post Company

[8] S. Marinkovic (2005), Section proposal related to Information Society for National Strategy of Serbia for accession to EU, Center for Internet development in Serbia

[9] National Information Technology and Internet Agency (2006), Internet portal of e-Government in Serbia

[10] The Ministry of Justice, The Republic of Serbia (2006), National Strategy for Judicial Reform

[11] B. Perovic etc. (2007), Ministry of Justice, Judicial Information System of Serbia

[12] The Ministry of Justice, The Republic of Serbia (2006), Annual Report for 2006

[13] The Ministry of Justice, The Republic of Serbia (2006), The Development Plan of Ministry of Justice for 2007-2001

[14] USAID (2007), Commercial Court Administration Strengthening Activity (CCASA) – Report

[15] The Ministry of Finance, The Republic of Serbia (2006), Information system in function of controlling public finance

[16] The Ministry of Finance, Custom Administration, Republic of Serbia (2006), The Modernization of the Serbian Customs Information Systems

[17] B. Dobrosavljevic, The Role of Information Technologies in E-Government - Serbian Business Registers Agency Experience, *YU Info 2007*

Growth of eGovernment Services in Macedonia (Online Sophistication of eGovernment Services)

Marjan Gusev, Dejan Spasov and Goce Armenski
University Sts Cyril and Methodius, Skopje
E-mail: marjan@on.net.mk, spasov.dejan@gmail.com, armenskig@gmail.com

Keywords: eGovernment, e-Business, benchmarking

Received: May 5, 2007

This study shows the enormous growth of eGovernment services in Republic of Macedonia for the period of 2004 up to 2007. It is shown that in comparison to the EUmember states and new member states Macedonia has also growing trend and will reach the objectives set in EU states, but at least 5 years later. Plenty of activities have been realized to enable this enormous growth, such as, establishing of National Strategy for Information Society (as possibility for donors to see where Macedonia would like to be and how it will reach the goal), realization of government contract with Microsoft for establishing basic eServices portal, and finally the most beneficial are the donor programs to support Macedonia, especially those by USAID, UNDP and EAR.

This growing tendency shows that now (2007) the Republic of Macedonia has already successfully realized the stage of establishing the infrastructure, and now is realizing the stage of establishing services. We can only comment that now (2007) EU states are in the stage of increasing the usage by establishing more content and interoperable functions. The comparison shows that the stage that Macedonia reaches in 2007 is comparable to the stage that most EU countries have reached in 2001/2002.

Povzetek: Povzetek opisuje uspešen razvoj e-uprave v Makedoniji v letih 2004 do 2007.

1 Introduction

The study [1] performed by the authors in April 2006 analyzed the current situation of eGovernment in Western Balkan Countries (WBC). As input the authors used eGovernment benchmarks and analyses of National Information Society (IS) Strategies. The existing EU eGovernment benchmarks are measured by Capgemini with methodology to measure *online sophistication of basic public online services* and *percentage of fully available online services* [2].

We evaluated the growth of eGovernment in Macedonia, comparing the progress in the area of eGovernment services with the EU countries. We also analyzed the respective National IS Strategy and conditions that enabled this growth, especially those parts which cover realization of eGovernment concepts. This study analyzes the action plans level of implementation and also present several issues given by Stability Pact eSEE initiative [3].

In this study we present the results of the third measurement on online sophistication of the basic public services in Republic of Macedonia (MK). The study was conducted in March 2007 with the purpose to compare the online sophistication in Republic of Macedonia with the reports for online sophistication in the EU countries from 2001 to 2006. In the report we closely monitor the online sophistication of the basic public services provided by the national government and two municipal authorities for the cities of Skopje and Veles.

In section 1 we will explain the methodology that we used in the survey. We monitor the same parameters that were monitored by Capgemini for eCroatia 2005 [4]. In section 2 we present our benchmarking results. We present the progress from 2004 to 2007, and in section 3 we describe in detail the conclusion and credits. The appendix lists situation with each service and the best practices in Republic of Macedonia.

2 Methodology

The 20 basic public services whose online sophistication level has been monitored are given in table 1. These services have been defined by the European Commission and monitored by Capgemini for EU and Croatia.

The service Social Security Benefits is made of the following sub-services: unemployment benefits, child allowances, medical costs, and student grants. The service Personal documents consists of: ID, Passport, and driver's license. Certificates are birth, marriage, and death certificates. There are few sub-services used by Capgemini for eCroatia that we also monitor for MK.

Each elementary service or sub-service is graded on a scale from 0 to 4. For example, social security benefits service encompasses unemployment benefits, child allowances, medical cost reimbursement, and student grants. Each of these four sub-services will be graded with 0, 1, 2, 3, or 4. Loosely speaking grade 0 is interpre-

	Citizens
1	*Income taxes*
2	*Job search*
3	*Social Security Benefits*
4	*Personal documents*
5	*Car registration*
6	*Building permission*
7	*Declaration to the police*
8	*Public libraries*
9	*Certificates*
10	*Enrollment in higher education*
11	*Announcement of moving*
12	*Health related services*

	Business
13	*Social contributions*
14	*Corporate tax*
15	*VAT*
16	*Registration of a new company*
17	*Submission of data to statistical offices*
18	*Customs declaration*
19	*Environment-related permits*
20	*Public procurement*

Table 1: The 20 basic services

ted as no information available on line; 1 is interpreted as relevant information available; 2 is interpreted as one way interaction; 3 as two way interaction; and 4 as transaction, or the service is fully available online.

European commission has defined maximum possible grade for each of the 20 services. Most services have maximum grade 4, but some, like Certificates and Job search, can have maximum grade 3. In order to provide better results for comparison, in this report we have used the Capgemini's maximum grade system, where job search has maximum grade 4.

The online sophistication for an elementary service is calculated in percents (%) as the ratio between the grade and the maximum attainable grade. The online sophistication of a basic service, like social security benefits, is the average of the four elementary services unemployment benefits, child allowances, medical cost reimbursement, and student grants. The final online sophistication level is the average of the sophistication of the 20 basic services.

At the end we should mention that each service can fall in one of the two target groups: citizens or business. In addition, there are four clusters that further divide the 20 basic services, named as income cluster, registration cluster, returns cluster, and permits cluster.

3 Results

Figure 1 lists all 20 public services together with their online sophistication score. Figure 2 show the progress for the basic services from 2004 to 2007.

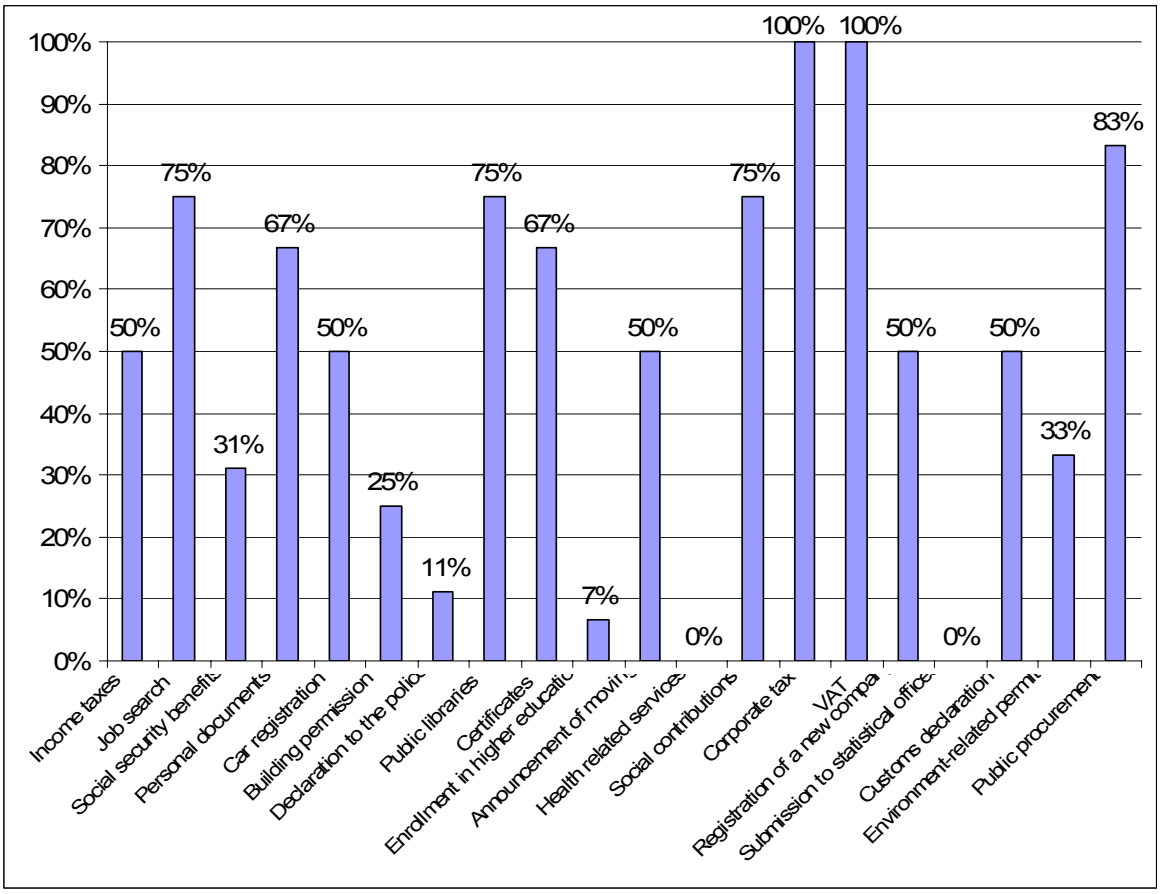


Figure 1: Online sophistication of the 20 basic services in Macedonia in March 2007.

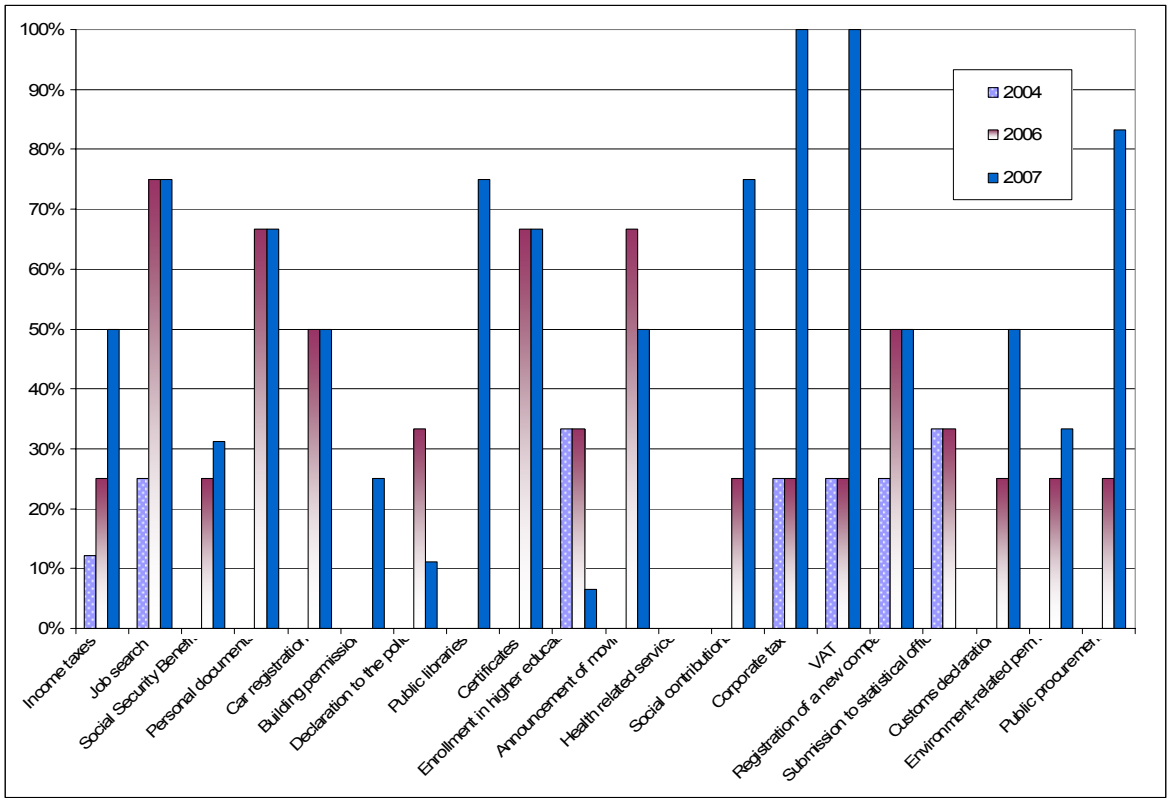


Figure 2: Progress of online sophistication of the 20 basic services in Macedonia for period of 2004-2007.

The online sophistication (OS) of Republic of Macedonia is computed as the average of the 20 basic services. The first measurement in 2004 has reported OS of 9%, in March 2006 MK has average OS of 32.75%, and in March 2007 the OS is 50%. For the last measurement we have used stricter grading criteria and we have included additional sub-services in the monitoring. For example for the sub-service "Enrollment in Higher Education" in year 2007 we use the following sub-services: Enrollment of new students, enrollment in higher educational year/ repetition of a year, registering for an exam, registering for an exam, certificate issuing,

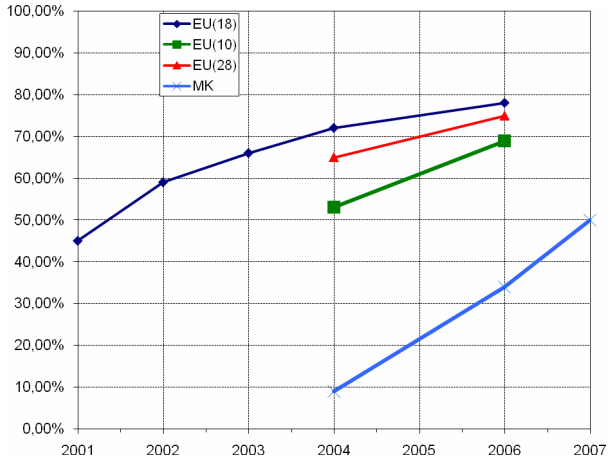


Figure 3: Online Sophistication comparison between Macedonia and EU.

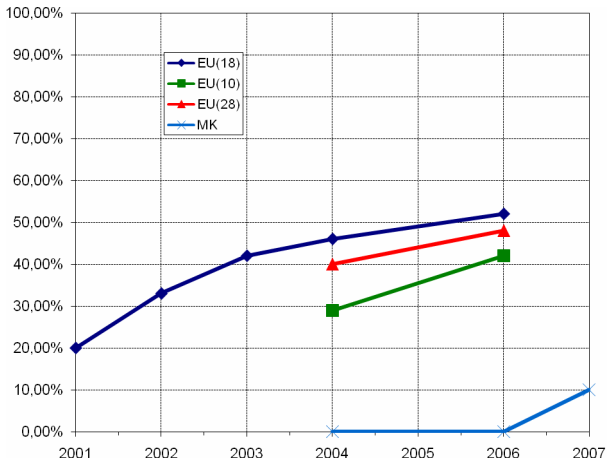


Figure 4: Percentage of fully available online services.

Figures 3 and 4 compare the OS and fully availability, accordingly, of MK with the OS of the EU. From Figure 3 and the report from Capgemini, we can conclude that Macedonia has OS level that is between Slovakia and Latvia. EU(28) refers to the all 28 states, EU(10) refers to the 10 new member states, and EU(18) represents the old member states plus Norway, Iceland, and Switzerland.

The first 12 services measure the online sophistication of the citizen's services, and the remaining 8 are for the business' services. Figure 5 shows the online sophistication growth for the citizen's services for

MK and EU countries. From both charts we can conclude that Macedonia has reached the EU's level in 2001, but not the level from 2002. With only 2 services fully available online, Macedonia has 25% fully availability sophistication. This means that Macedonia hasn't reached the level of EU in 2001.

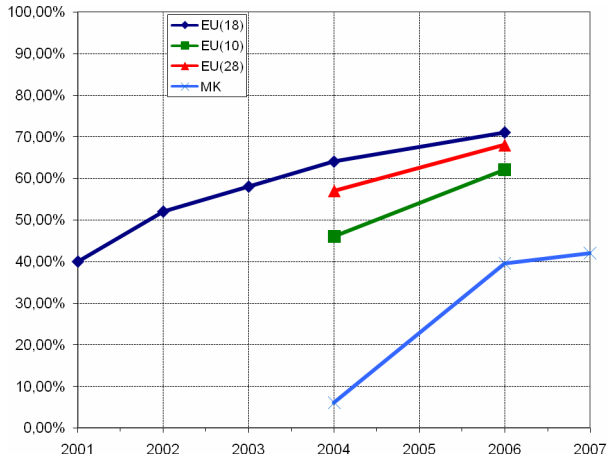


Figure 5: Online sophistication of the citizen's services.

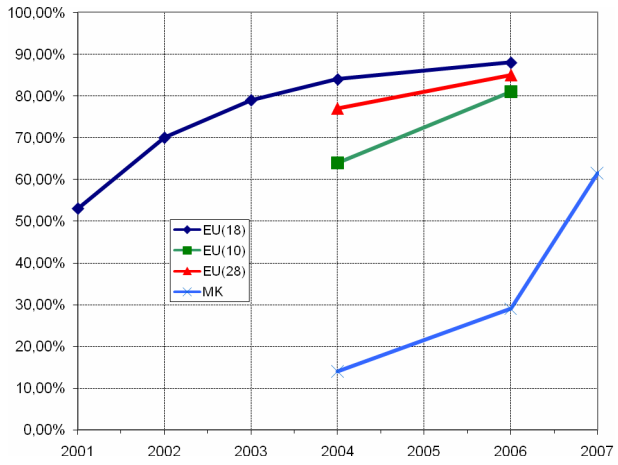


Figure 6: Online sophistication for the business' services.

The next figures show the results from the clusters.

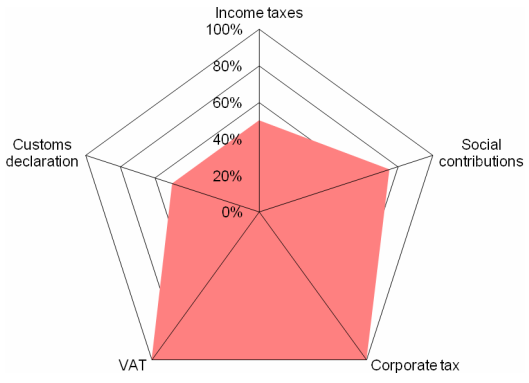


Figure 7: The income generating cluster.

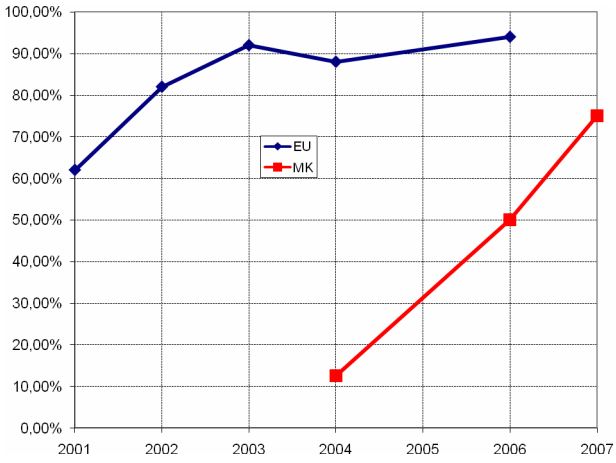


Figure 8: The progress in the income cluster

The income generating cluster is made of services where finance flow from citizens and business to the government institutions. Figures 7 and 8 show the progress. In 2006 EU(28) had a score for OS of 94%, while MK in '07 has a score 75%. The new member states in 2005 had online sophistication of 74%.

Registration cluster encompasses services for mandatory information storage. Figures 9 and 10 show the progress of the registration cluster.

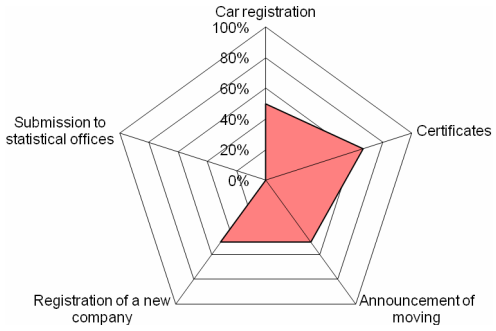


Figure 9: The registration cluster.

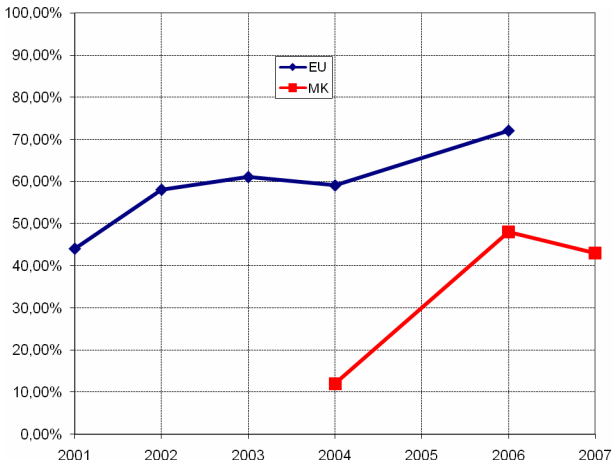


Figure 10: The progress of the registration cluster.

Returns cluster is for services given to citizens and business in return for taxes and contributions. Figures 11 and 12 show the results.

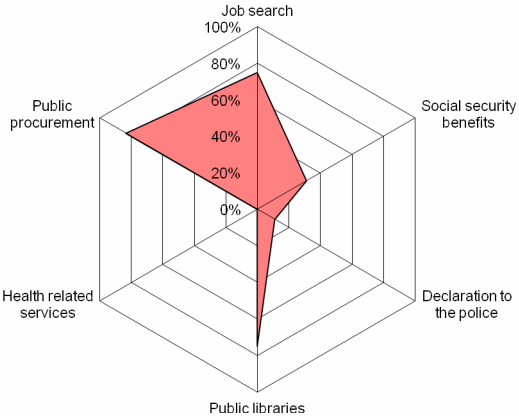


Figure 11: The returns cluster.

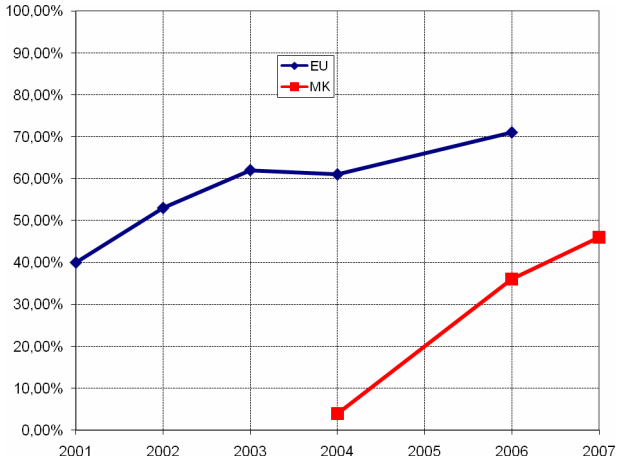


Figure 12: The progress of the returns cluster.

Permits and Licenses cluster represents documents provided by the government. This service is largely decentralized, so the improvements will be the hardest. Figures 13 and 14 show the progress and the results.

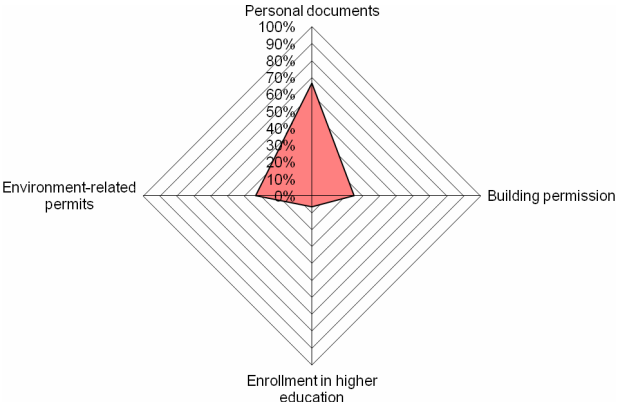


Figure 13: The permits cluster.

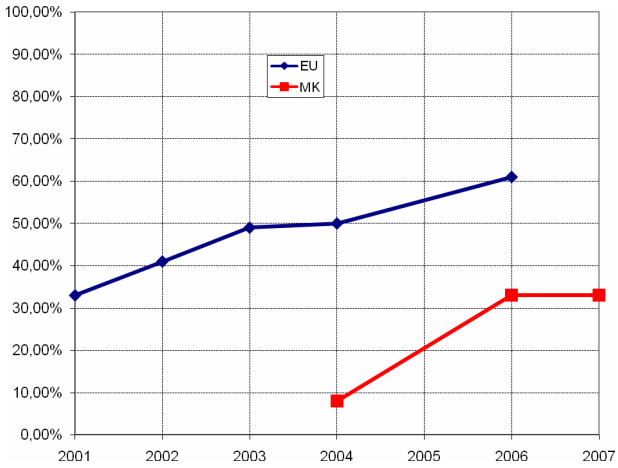


Figure 14: The progress of the permits cluster.

From the figures we can see that the income cluster has increased its sophistication level for 200%, and the returns cluster for almost 100%. The situation in the other two clusters is slowly improving, although this improvement is hard to see because of the stricter criteria that we use this year.

On the other hand, if we observe each of the 20 basic services we will spot the very low sophistication for the following services: declaration to police, enrollment in higher education, health-related services, and submission data to statistical office. If we set as a short term goal to make each of the 20 services with minimum grade 2, we will have online sophistication of around 65%, which is very close to the new EU countries average.

4 Conclusion

An enormous growing trend has been achieved in absolute values for on-line sophistication of eGovernment services in Macedonia for the period 2004-2007. The value of 7,89% was reached in April 2004, the value of 32,89% in April 2006 and the value of **50% in March 2007**. This linear growth rate of eGovernment services shows tendency as realized in other EU states. According to this growth it is likely to expect values of average EU states online sophistication of 75% in 3-4 years.

A percentage of fully available online services is only 10%, showing growth from zero value in 2006. Assuming the same growing tendency we can conclude that Macedonia will reach the average EU value of 50% in 4-5 years.

Although the grow rate is enormous we can conclude that Macedonia has the same online sophistication as average value of EE member states in 2001. This can only lead to a conclusion that Macedonia is 5-6 years behind.

The big difference between the situation (and high grades) in the business eGovernment services and the services in the citizen's sector is mainly due to donor

programs that pushed development of businesses and not citizens sector.

Credits

Three main reasons can be extracted for enabling such enormous growth. The first reason is the realization of National Strategy for Information Society and clear vision where Macedonia would like to be in near future and how it will be achieved. The second reason is enormous help of donors, especially USAID and the eGov project, mainly due to the USAID's contribution to eTaxes and eProcurements. Other donors are UNDP mainly in the municipalities and taxation and EAR in the business sector. The third factor for this enormous growth is the willingness of the government to support activities for eGovernment and realize several projects, such as contract with Microsoft and other companies for realization of eServices, eGoverment, eParliament and other software packages. We can also extract the contribution of the Committee for Information Technology as leader for all these activities.

5 Comparison and future

EU has realized two development stages in the past (eEurope 2002 and eEurope2005) and is now realizing the i2010 initiative. The first stage was establishing infrastructure and the second is establishing services. The main motivation for the eGovernment benchmarking was development of a tool to measure the level of established services and now there is a tendency to establish new tool since most of the countries have reached satisfactory level of established infrastructure and eGovernment services. The new tool should express how much is the usage level, or how much the citizens and businesses use appropriate services. This can be easily seen from a comparison chart that expresses demand and usage.

The holistic measurement model shows existence of 5 stages for development: setting (equivalent to establishing infrastructure); delivery (equivalent to establishing services); change (equivalent to reorganization to enable one stop shop paradigm); use (equivalent to usage) and impact (the final stage). It is obvious that most of EU states are now in stage 3 and 4, while Macedonia is in transition from stage 1 to 2.

Therefore a new aggressive approach can make Macedonia to reach the EU stages sooner by setting a clear vision and strategy how to reach these objectives.

References

[1] Gusev, Armenski (April 2006): Gap Analysis of eGovernment in Western Balkans, www.metamorphosis.org.mk, 12.05.2006

[2] Capgemini. (June 2006). Online Availability of Public Services: How Is Europe Progressing? Web Based Survey on Electronic Public Services Report of the 6th Measurement June 2006. Capgemini, (http://ec.europa.eu/information_society/eeurope/i2010/docs/benchmarking/online_availability_2006.pdf)

[3] Stability Pact for South Eastern Europe, eSEE Working Group (www.stabilitypact.org/e-see).

[4] Capgemini, Study on online availability of public services in 2005 for eGovernment in Croatia; http://www.e-croatia.hr/repozitorij/dokumenti/downloads/Online_Availability_2005.pdf

Appendix: The 20 Basic Public Services

Here we explain each of the 20 basic public services and how we set the grade for Republic of Macedonia.

1. Income taxes. Online sophistication of 50%

The Central Taxation agency is the service provider for this service. On their web there is a possibility to download form to start the procedure. Their web page is www.ujp.gov.mk

2. Job search. Online sophistication of 75%

According to Capgemini's methodology, this service can have max grade 4 if the service provider offers the possibility of an electronic supply of pre-selected jobs related to the given profile of the job searcher.

Employment Service Agency of the Republic of Macedonia offers the possibility for online search of the database with job offerings. However we were unable to locate the URLs for the 30 regional employment centers.

The URL for the service provider is: http://www.zvrm.gov.mk

3. Social Security Benefits. Online sophistication of 31.25%

There are four services in this category defined by the EU Commission

a. *Unemployment benefits*; graded with 1

b. *Child allowances*; graded with 2

c. *Medical costs reimbursement*; graded with 1

d. *Student grants*; graded with 1

Different service providers are involved in providing social security benefits. Ministry of Labor and Social Policy is responsible for Unemployment benefits and Child allowances. Only child allowance have grade 2. Information for this service can be obtained through

www.uslugi.gov.mk. The description of the other services in this group can be found:
- for Unemployment at: www.mtsp.gov.mk
- for Medical costs at: www.fzo.org.mk
- for Student grants at: http://www.mon.gov.mk/

4. Personal documents. Online sophistication of 66.6%

In Macedonia personal documents are considered to be Passport, ID, and driver's license. Ministry of interior is the service provider. Citizens can obtain information and download forms at www.uslugi.gov.mk. The EU Commission has defined max grade for this service to be 3.

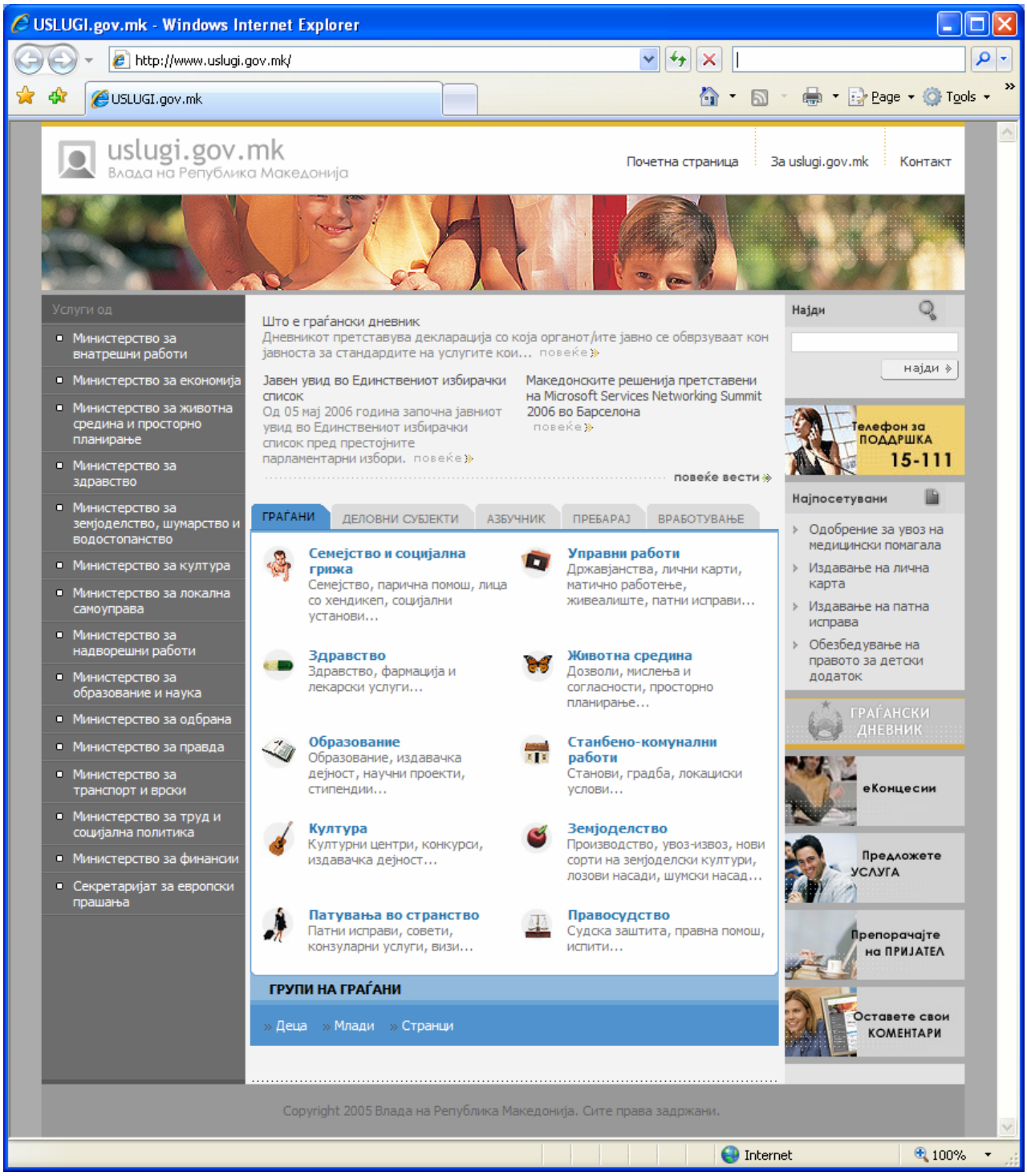


Figure 15: Best practices: www.uslugi.gov.mk

5. Car registration. Online sophistication of 50%

Same service provider as in 4. The maximum attainable grade is 4.

6. Building permission. Online sophistication of 25%

The information on how to start the procedure for obtaining building permission can be found at http://www.e-skopje.gov.mk/. The information for obtaining evidence of property can be found at www.katastar.gov.mk.

7. Declaration to police. Online sophistication of 11%

Ministry of Interior has information on their web site regarding crime reporting and crime prevention procedures. However this parameter measures the online sophistication of the local police stations too. We were unable to locate any URL for the local police authorities.

8. Public libraries. Online sophistication of 75%.

According to Capgemini's methodology, this service can have max grade 4 if the service provider offers the possibility to search for specific title and to make electronic reservation, or to obtain electronic copy.

Macedonia as a part of the co-operative online bibliographic system COBIS.Net has publicly accessible website with the possibility for online search. There are 21 libraries in Macedonia included in this project.

The URL for the service provider is: http://vbmk.nubsk.edu.mk/cobiss/

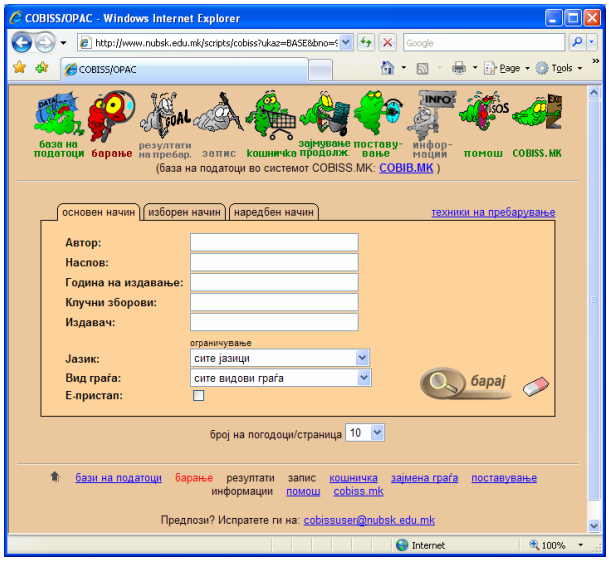


Figure 16: Best practices: Public Libraries

9. Certificates Online sophistication 66.6%

Same service provider as in 4. Maximum grade for this service is 3.

10. Enrollment in higher education. 6.6% of online sophistication

This year we are measuring the following 5 subservices:

1) *Enrolment in higher educational institution*; graded with 1
2) *Enrolment into higher year/ Repetition of the same year*; graded with 0
3) *Registering for an exam* ; graded with 0
4) *Certification issuing*; graded with 0;
5) *Student grants/ scholarships*; graded with 0

Except for the copy of the public announcement for enrolment in higher educational institution published on the web, we were unable to find any information regarding the other four parameters. We take into account only the universities that receive government's funding.

11. Announcement of moving. Online sophistication of 50%

The service provider is the same as in 4. The announcement of moving refers to change of the address within the country. In republic of Macedonia there are 2 elementary sub-services that can be classified under this category:

- Change of the address in a personal document; graded with 1
- Change of the city of residence; graded with 2

12. Health related services. Online sophistication of 0%

We were unable to locate any information on the web regarding interactive advices for hospitals. The majority of hospitals don't have even a web page.

Health related services include:

- Interactive consultation of available services;
- Interactive appointments

13. Social Contributions. Online sophistication of 75%

Pension and Disability Fund of Macedonia is responsible for this service. The provider has a web service with the possibility to completely treat the declaration of social contributions for

employees via website. Although this service deserves grade 4, PIOM has been graded with grade 2 for treating the new employments. The final grade is the average between these two grades.
The URL is: http://www.piom.com.mk/

14. Corporate tax Online sophistication of 100%

Paperless taxes were part of the USAID's eGov initiative. The service provider is the Central Taxation Agency. The system is located at http://etax.ujp.gov.mk/

15. VAT Online sophistication of 100%

Same as 14

Figure 17: Best practices: Central Taxation Agency

16. Registration of a new company. Online sophistication of 50%

Central registry is responsible for this service. On their web you can download forms to register a new company.
The URL is: www.crm.org.mk/

17. Submission of data to statistical office. 0% of online sophistication

The republic's statistical office has a web page with plethora of information regarding the statistical data for Republic of Macedonia. Regarding the submission of data, we were able to find some telephone numbers. This year we have decided to treat this as no relevant information. The main URL is www.stat.gov.mk

18. Customs declaration. Online sophistication of 50%

Customs office has publicly accessible website with the possibility to download forms. The service provider is located at www.customs.gov.mk

19. Environment-related permits online sophistication 33.33%

Ministry of Environment and Physical Planning offers the possibility to download

some forms through www.uslugi.gov.mk. Local governments, like Municipality of Veles on their website (www.veles.gov.mk) has the information on how to start the procedure for obtaining the Integrated Pollution Prevention and Control permit.

20. Public procurement Online sophistication of 83.33%

The USAID's eGov project in cooperation with the Public Procurement Bureau has developed e-Procurement system for on-line conducting public procurements. The system is located at https://e-nabavki.gov.mk and has been used by the local authorities. The national government has online sophistication of 50%.

Modeling eGovernment processes with UMM

Philipp Liegl and Rainer Schuster
Studio Digital Memory Engineering, Research Studios Austria
Thurngasse 8/20, 1090 Vienna
{pliegl, rschuster}@researchstudio.at, http://dme.researchstudio.at

Marco Zapletal
Electronic Commerce Group, Vienna University of Technology
Favoritenstr. 9-11/188, 1040 Vienna
marco@ec.tuwien.ac.at, http://www.ec.tuwien.ac.at

Christian Huemer
Business Informatics Group, Vienna University of Technology
Favoritenstr. 9-11/188, 1040 Vienna
huemer@big.tuwien.ac.at, http://www.big.tuwien.ac.at

Birgit Hofreiter
Faculty of Information Technology, University of Technology Sydney
PO Box 123, Broadway NSW 2007, Australia
birgith@its.uts.edu.au, http://www.it.uts.edu.au/

Robert Mosser
Inst. of Distributed and Multimedia Systems, University of Vienna
Liebiggasse 3/3-4, 1010 Vienna
robert.mosser@univie.ac.at, http://www.cs.univie.ac.at

Keywords: inter-organizational business processes, UMM, UML, eGovernment

Received: July 31, 2007

The United Nation's Center for Trade Facilitation and Electronic Business (UN/CEFACT) is a standardization body known for its work on UN/EDIFACT and ebXML. One of its most recent developments is UN/CEFACT's Modeling Methodology (UMM). The UMM standard is used to model inter-organizational business processes in the B2B domain. With the increasing availability of electronic governmental services over networks, the frontier between B2B and B2G/G2G disappears. Today one expects a governmental institution to react like any other business partner. Therefore also governments now face the interoperability and compatibility issues as regular businesses do. In order to allow two governmental institutions to collaborate, a methodology uniquely depicting the inter-organizational process from a global perspective is needed. In this paper we propose to use UN/CEFACT's Modeling Methodology in the eGovernment domain. UMM allows the definition of a global choreography which is then being used to derive local orchestrations for each business partner. Such orchestrations can then be used by enterprise applications in a service oriented context. As an example, a real-world scenario from the waste transport domain within the European Union will be shown. Furthermore the possible integration in the context of the We-Go project [3] is examined.

Povzetek: Predstavljeno je modeliranje procesov e-uprave z UMM.

1 Introduction

The multitude of eGovernment initiatives and efforts which have been made in the last few years have shown the emerging significance of the domain. With more and more governmental institutions being present on the Internet, other businesses expect a governmental institution to react like a regular business does e.g. offering key business functionalities via services accessible over the Internet. When fulfilling these expectations governmental institutions face the same problems as regular businesses. Therefore the distinction commonly made between government-to-government (G2G), business-to-business (B2B), and business-to-government (B2G) is not appropriate any more.

In the context of eGovernment a distinction is made between informational services, communicational services and transactional services. Information services provide content to citizens and enterprises in an unidirectional way. Communication services are providing information content

too but in a bidirectional way, hence allowing feedback. Most of the eGovernment services in use today are either informational or communicational services. Whereas these two service types do not impose real technological challenges, transactional services are more difficult to implement. A transactional service allows enterprises and citizens to interact with governmental institutions in order to execute official functions e.g. announce a waste transport from country *A* to country *B*.

Business interactions which before were done face-to-face with a government official are now conducted electronically over a network. The implementation of such a transactional service becomes more difficult if many enterprises are actively using the service in an automated manner. In particular larger enterprises will use the interfaces provided by the government and invoke them through their own enterprise system. If such a process is conducted in an automated manner between two independent electronic systems, the business process itself and in particular the order in which the business messages are exchanged, must be well defined. Figure 1 shows an example of transactional services in the eGovernment domain. In order to perform transactional services between two business partners a set of alignments is necessary.

In the first step the legal requirements must be set and well-defined. Before an eGovernment transaction can be executed the necessary policies and guidelines must be set. Policy alignment means that the involved parties must act in a specific legal area. Given that the parties operate in a well-defined and safe legal framework, there still may be open legal space which must be filled by contracts between the parties. Therefore, a contractual alignment between the parties is required in order to close any legal gap.

Directly influenced by the legal settings is the organizational layer. This layer focuses on the alignment of the shared business logic between the two participating parties. Once the involved business partners have agreed upon a global perspective, the process can be executed. Hence the "contract" defined on the organizational layer governs the technical execution e.g. the exchange order of business messages (process alignment) and what business messages are exchanged (information alignment). In terms of business alignment the business layer defines a global choreography to which the two business partners must comply. The global choreography is then used to derive the local orchestrations for each business partner. This guarantees that both local orchestrations comply to the global choreography and are thus complementary.

The choreography defined on the organizational layer must be executed between the two involved information systems. In other words, a workflow of business document exchanges must be implemented. In order to keep the business applications of the involved business partners as flexible as possible to the inter-organizational business process, the use of a business service interface is suggested. A business service interface as described in figure 1 forms a sort of proxy layer between the own business application and the business application of the business partner. The business service interface uses a declarative approach to define the execution order of the business process (e.g. BPEL) and can therefore be quickly adjusted to any changes in the underlying business process.

In order to help governmental institutions to implement transactional services we propose to use UN/CEFACT's Modeling Methodology. The United Nations Center for Trade Facilitation and Electronic Business (UN/CEFACT), known for its standardization work in the field of UN/EDIFACT [23] and ebXML [16], has developed the so called UMM standard. UMM stands for UN/CEFACT's Modeling Methodology and enables to capture business requirements independent of the underlying implementation technology such as Web Services or ebXML. In regard to figure 1 UMM provides a formal mechanism to define the requirements for the organizational and technical layer.

The remainder of this paper is structured as follows: Section 2 explains the UMM standard in detail by using an example from the waste management domain. In section 3 we introduce our UMM modeling tool called UMM Add-In and its support features for the UMM standard. The integration of the UMM standard in the roll-out of the We-Go project will be the subject of section 4. Section 5 gives an overview about other research efforts in the field of inter-organizational business processes. The paper concludes with an outlook and future research issues.

2 UN/CEFACT's Modeling Methodology

UN/CEFACT's Modeling Methodology (UMM) is a UML based methodology for capturing the requirements in an inter-organizational business process. It is independent of the underlying transfer syntax. The overall goal of the UMM methodology is to create a global choreography of the business process. If two business partners interacting with each other both define their own choreography for the business process, the resulting choreographies are unlikely to match. UMM pursues a top down approach by first defining the global choreography from which the local choreographies are derived. Hence it is ensured, that both choreographies are complementary.

UMM is built upon the UML meta model and defined as a UML profile [25] e.g. a set of stereotypes, tagged values and OCL constraints. Its current version is 1.0 based on UML 1.4 [17]. The predecessor based on UML 2.0 is currently under development.

In the following section we will outline a real-world UMM example from the eGovernment domain in the European Union.

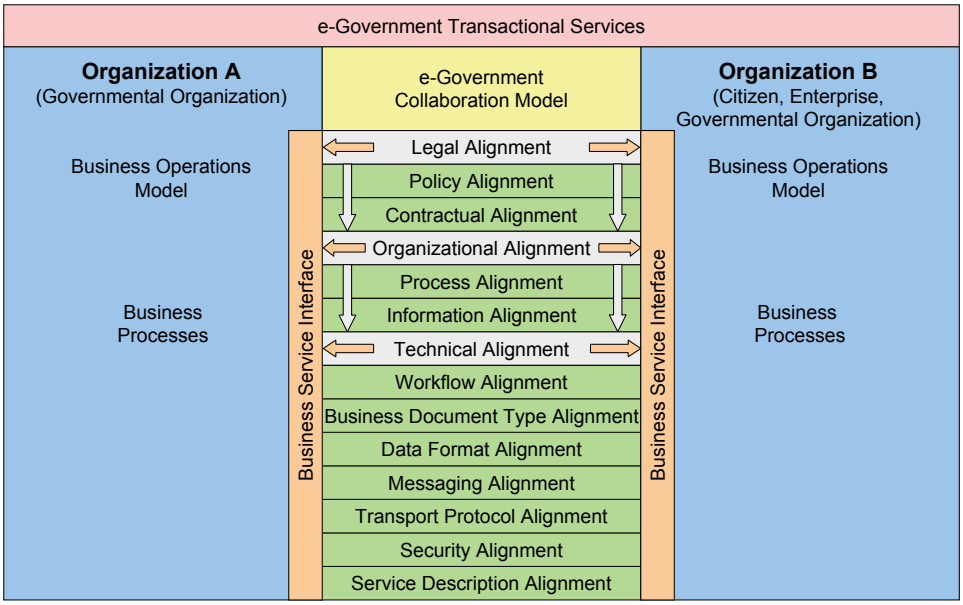


Figure 1: Transactional services in the eGovernment domain

2.1 A UMM example from the eGovernment domain

In this section we briefly describe the steps of UMM and the resulting artifacts. For a better understanding we walk through the UMM by the means of a simple example from the waste transport domain as it is used in the European Union today. A waste transport taking place between two countries in the European Union will be analyzed. The transport of goods and persons within the EU is not subject to strict regulations any more. However the transport of waste between two European countries underlies regulations accompanied by a multitude of forms and administrative documents. One major goal of the European Union is to facilitate the waste transport by supporting the transport process with information technology means and to decrease the amount of paper documents involved.

The relevant artifacts of our example are depicted in figure 2. On the left hand side of this figure we see the structure of our *waste management* model. A UMM *business collaboration model* comprises three main views: the *business domain view* (BDV), the *business requirements view* (BRV), and the *business transaction view* (BTV). The three top level packages of any UMM model are always stereotyped in accordance to these views. For relevant artifacts in the tree-view on the left hand side of figure 2 the according diagrams are shown on the right hand side of the figure.

The BDV is used to gather existing knowledge from stakeholders and business domain experts. In interviews the business process analyst tries to get a basic understanding of the business processes in the domain. The use case descriptions of a business process are on a rather high level. One or more *business partner types* participate in a business process and zero or more stakeholders have an *interest in* dependency with the process. The BDV results in a map of business processes, i.e. the business processes are classified. Thus, the BDV package includes *business area* sub-packages. UN/CEFACT suggests to use *business areas* according to the Common Business Process Catalog [24]. Each *business area* consists of *process area* packages that correspond to the Open-edi phases (planning, identification, negotiation, actualization, and post-actualization) [10]. In our *waste management* example the relevant *business areas* are *logistics* and *regulation*. *Logistics* covers at least the *process areas* of *actualization* and *post-actualization*. The *business area regulation* only contains the *process area actualization*. We do not want to detail here all the business processes that may be important to the domain experts and stakeholders in these areas. Due to size restrictions we do not present the structure of *business areas* and *process areas* in the *business domain view* of figure 2.

The BRV consists of a number of different sub-views. The *business process view* (1) and the *business entity view* (2) are both very project specific. The *business process view* gives an overview about the business processes, their activities and resulting effects, and the business partners executing them. The activity graph of a business process may describe a single partners process, but may also detail a multi-party choreography. The business process analyst tries to discover interface tasks creating/changing business entities that are shared between business partners and thus, require communication with a business partner. Discovery of interface task is more important in this step than modeling an exact control flow of activities. In our example we detail a multi-party business process for an *end-to-end waste transport* (1). The four parties involved in the process are *exporter*, *export authority*, *import authority* and *importer*. Due to space limitations the *exporter* and the *im-*

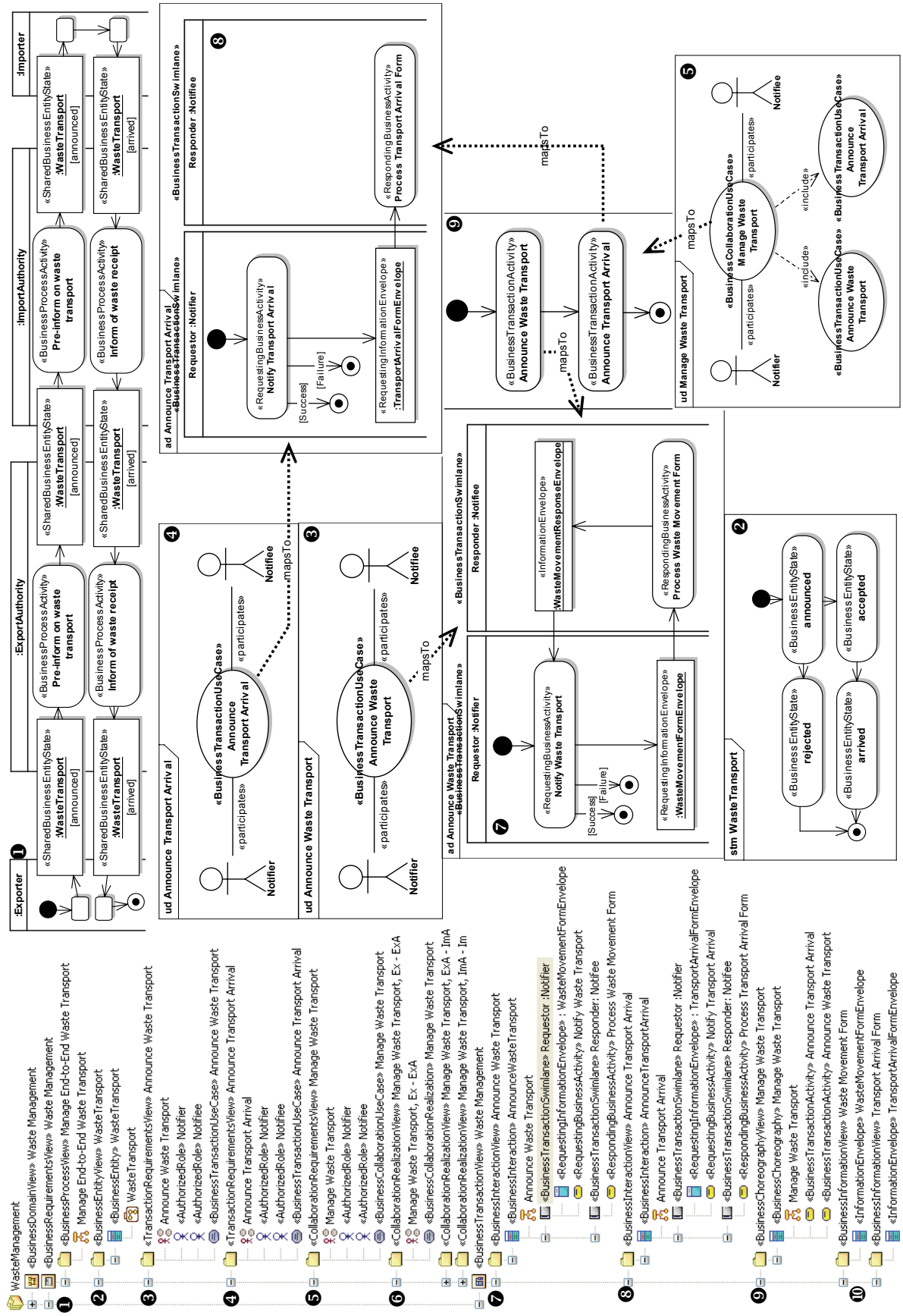


Figure 2: UMM Overview

porter have been simplified. The *exporter* pre-informs the *export authority* about an upcoming *waste transport*. The *export authority* in turn informs the *import authority* of the country the *waste transport* is imported to. The *import authority* then informs the *importer*. After the *waste transport* has been conducted the *importer* informs the *import authority* about the arrived *waste transport*. The *import authority* then informs the *export authority* of the country the *waste transport* originated from. Finally the *export authority* notifies the *exporter* about the successful *waste transport*.

The information exchanged between business partners is about the business entities *waste transport*. A *waste transport* object is created with the state *announced*. Later it is set either to the state *rejected* or *accepted*. Finally, after the waste transport took place the *waste transport* state is set to *arrived*. These so-called *shared business entity* states must be in accordance with the *business entity lifecycle* of *waste transport*. The lifecycle is defined in the state chart of the *business entity view* (2). Due to space limitations the process view in (1) has been simplified and the options for setting the *waste transport* to *accepted* or *rejected* have been left out.

Since the same tasks take place between a different pair of *business partner types*, it is not appropriate to describe these tasks for each pair again and again. Instead, these tasks are defined between *authorized roles*. A *transaction requirements view* defines the *business transaction use case* for a certain task and binds the two *authorized roles* involved. The *authorized roles* are defined in the exact context of the *business transaction use case*. In our example we have two *transaction requirements views*: *announce waste transport* (3), and *announce transport arrival* (4). By coincidence the *authorized roles* in *announce waste transport* and *announce transport arrival* have the same name, namely *notifier* and *notifiee*.

The *collaboration requirements view* includes a *business collaboration use case*. The *business collaboration use case* aggregates *business transaction use cases* or nested *business collaboration use cases*. This is manifested by *include* associations. In our example the *business collaboration use case manage waste transport* (5) includes the *business transaction use cases announce waste transport* (3) and *announce transport arrival* (4). Furthermore, the *authorized roles* participating in the *business collaboration use case* must be defined within the context and namespace of the *collaboration requirements view*. We call the roles *notifier* and *notifiee*. Again the homonymous names are coincidental. The *notifier* is the one who initiates the *manage waste transport* and the *notifiee* is the one who reacts on it. The *authorized role notifier* in *announce waste transport arrival* (4) is not the same as the one in *manage waste transport* (5). This means, we have two *authorized roles notifier*, each defined in the namespace of its view. *Maps to* dependencies are used to define which *authorized role* of a *business collaboration use case* plays which role in an *included business transaction use case* (or nested *business collaboration use case*). In our example the *notifier* of *manage waste transport* (5) plays the *notifier* of *announce waste transport* (3) but also the *notifier* in *announce transport arrival* (4). The *notifiee* of *manage waste transport* (5) plays the *notifiee* of *announce waste transport* (3), but also the *notifiee* of *announce transport arrival* (4).

A *business collaboration use case* may have many *business collaboration realizations* that define which *business partners* play which *authorized roles*. Due to space limitations the concept of *business collaboration realizations* is not further elaborated here.

The BTV builds upon the BRV and defines a global choreography of information exchanges and the document structure of these exchanges. The choreography described in the requirements of a *business transaction use case* is represented in exactly one activity graph of a *business transaction*. A *maps to* dependency between them allows traceability between the requirements and the *business transaction*, which is defined in a *business interaction view*. In our example, the *announce transport arrival* requirements (4) are mapped to a corresponding choreography (8). The same mapping is made for the *announce waste transport* requirements (3+7).

A *business transaction* is characterized as follows: If an *authorized role* recognizes an event that changes the state of a *business entity*, it initiates a *business transaction* to synchronize with the collaborating *authorized role*. A *business transaction* is an atomic unit that leads to a synchronized state in both information systems. We distinguish one-way and two-way *business transactions*: In the former case, the initiating *authorized role* reports an already effective and irreversible state change that the reacting *authorized role* has to accept. In the other case, the initiating partner sets the *business entity/ies* into an interim state and the final state is decided by the reacting *authorized role*. It is a two-way transaction, because *business information* flows from the initiator to the responder to set the interim state and backwards to set the final and irreversible state change. Irreversible means that returning to an original state requires compensation by another *business transaction*.

Owing to this strict definition, a UMM *business transaction* follows always the same pattern: A *business transaction* is performed between two *authorized roles* that are already known from the *business transaction use case* and that are assigned to exactly one *business transaction swimlane* each. Each *authorized role* performs exactly one activity. An *object flow* between the *requesting* and the *responding business activity* is mandatory. An *object flow* in the reverse direction is optional. In our example the *business transactions announce waste transport* (7) is a two-way transaction whereas *announce transport arrival* (8) is a one-way transaction. Sending the *waste movement form envelope* (7) sets the interim state *arrived* of the *business entity waste transport*. The reply in the *announce waste transport* sets it to the state *accepted* or *rejected*. Finally the notification in *announce transport arrival* (8) sets the final state of the *business entity waste transport* to *arrived*.

The requirements described in a *business collaboration use case* are choreographed in the activity graph of a *business collaboration protocol*, which is defined in a *business choreography view*. This one-to-one relationship is denoted by another *maps to* dependency. In our example, the *manage waste transport* requirements (5) are mapped to the *business collaboration protocol* shown in (9). A *business collaboration protocol* choreographs a set of *business transaction activities* and/or *business collaboration activities*. A *business transaction activity* is refined by the activity graph of a *business transaction*. In our example, the *business collaboration protocol* of *manage waste transport* (9) consists of two *business transaction activities*: *announce waste transport* and *announce transport arrival*. Each of them is refined by its own *business transaction* (7,8). *Maps to* dependencies keep track of this refinement. *Business collaboration activities* - which are not used in our example - are refined by a nested *business collaboration protocol*.

The *business information views* are used to define the structure of *business documents* exchanged in *business transactions*. Each of the four *information envelopes* exchanged in the two *business transactions* lead to a *business information view* describing the envelope's document structure. Due to space limitations only two *information envelopes* are shown namely *waste movement form envelope* and *transport arrival form envelope*. An *information envelope* consists of a *header* and a *body*. The *header* carries auxiliary information and the *body* holds the actual business document exchanged during the transaction. UMM itself does not prescribe a particular format for the exchanged business document. However the use of so called core components [26][28] is suggested.

3 A tool support for UMM

Since UMM is defined as a UML profile, a business analyst may use any UML tool to model UMM business collaboration models. As we outlined in the beginning, UMM artifacts are based on a specific subset of UML to capture complex business collaborations. Between these artifacts, a number of dependencies and constraints exist. If a regular UML tool is used for UMM, these rules are not enforced. Furthermore, a modeler is not prevented from using modeling elements that are not part of UMM's meta model. However, valid models are required for at least two reasons. Firstly, if a business process model is shared between partners, it has to be formally correct in order to ensure an unambiguous communication in terms of the modeled business domain. Secondly, a UMM model allows the generation of machine-interpretable process and business document specifications to configure B2B information systems. Obviously, deriving artifacts following a model driven approach works only for valid models. Consequently, a modeling tool considering UMM-specifics is needed. Instead of developing a new modeling tool from scratch, we decided to build our UMM tool on top of the commercial UML tool Enterprise Architect. The UMM Add-In utilizes Enterprise Architect's UML functionality and extends it by implementing the UMM-specifics. The UMM Add-in supports the business analyst by the following main features:

1. UMM-specific toolbar
2. UMM requirements engineering support
3. Semi-automatic generation of UMM artifacts
4. Validation of the UMM model
5. Generating process specifications for B2B information systems
6. Core Component support

3.1 UMM specific toolbar

In order to create a UMM model it is convenient to drag and drop UMM stereotypes from a toolbar onto the modeling canvas. Thus, the stereotypes as defined in the UML profile for UMM are integrated into Enterprise Architect and provided in a toolbar. The toolbar itself is organized in sections that correspond to the UMM views. This helps the user applying the right stereotypes in a certain UMM step. In order to prevent the user from using UML elements not part of the UMM subset, the toolbar is restricted to the UMM stereotypes and UML standard elements required by UMM.

3.2 UMM requirements engineering support

Requirements engineering is not only important for software development processes, but is also an critical task in business process modeling. Usually a lot of different people with different backgrounds are involved in the development process of a business model. In order to close this gap of various levels of business knowledge and expertise, the requirements of a business model must be documented in natural language. The outcome of this process is a common understanding of each part of the business scenario. Usually the business knowledge is collected during interviews between the business analysts and the business domain experts.

Especially in UMM the requirements engineering plays a major role. UN/CEFACT recommends using a set of so-called worksheets in order to capture the gathered business information. A UMM worksheet has its pre-defined structure containing standardized elements and is assigned to a specific part in the business model. Traditionally, worksheet templates have been filled out using a word processor. This leads to a great number of external paper documents. Since most of the worksheet information is captured later on in tagged values, this result in duplication of efforts and a danger of inconsistency. Thus the worksheet editor of

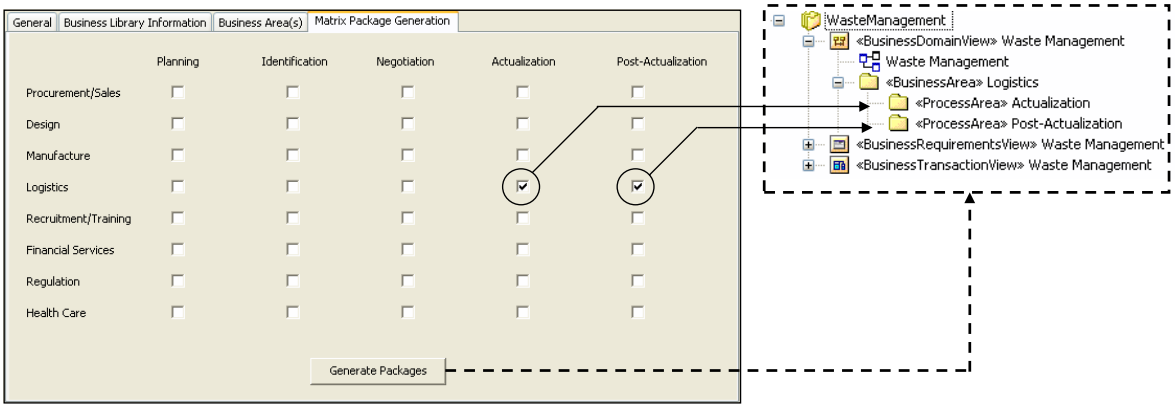


Figure 3: Semi-automatic generation of BDV's initial package structure

the UMM Add-In integrates the requirements engineering into the UMM model. Once the business domain knowledge is acquired, the information is directly stored in the corresponding tagged value of the assigned UMM artifact.

The worksheet editor offers an interactive form including all elements specified by UN/CEFACT. Additionally self-defined worksheets may be created using an XML based worksheet definition language (WDL). Each UMM artifact, which needs to be documented, has an XML based property file, in order to specify the layout, the structure and the content of the worksheet. This guarantees a flexible adaptation of worksheets to special business needs and to changes due to an update of the UMM meta model. In order to ensure traceability, some artifacts of a UMM model are interlinked - e.g. the *business transaction* in the *business transaction view* has a dependency to the corresponding *business transaction use case* in the *business requirements view*. Thus, business requirements information must not differ between such UMM worksheets. In order to avoid inconsistency the worksheet editor provides the modeler with business information already entered to prior worksheets having semantically the same content.

The communication between the business domain expert and the business analysts is usually based on some model reports. Thus, such reports must be generated from the model. The modeler is able to generate document reports using the export functionality of the worksheet editor. The user either exports the gathered information into Microsoft Word, or publishes HTML pages of the business model requirements. Furthermore, the whole documentation can be stored in an XML based format for backup purposes as the XML files can be re-imported into the model.

3.3 Semi automatic generation of UMM artifacts

In the section above we have demonstrated how the worksheet editor is used for capturing business requirements of a UMM model. The worksheet editor offers an added value by generating UMM artifacts automatically from the captured requirement information. Thereby, the modeler is released from routine tasks. Instead of modeling recurring UMM patterns manually the UMM Add-In creates such artifacts automatically using the input of the worksheets the modeler already entered.

In figure 3 such a workflow is depicted. The worksheet of the *business domain view* captures information about *business areas* and *process areas*. Since UMM recommends using a set of hierarchically pre-defined *business* and *process areas* according to the Common Business Process Catalog (CBPC) [24], the package structure of the BDV differs only by the modeler's choice of package constellations. The structure of *business areas* and *process areas* - which may follow a matrix - is transformed into a tree-structure of packages. Instead of manually creating this tree-structure, the modeler simply checks the relevant cells in the pre-defined matrix and the package structure is created. In our *waste management* example the *business area logistics* together with its *process areas actualization* and *post-actualization* have been generated automatically using this matrix.

Another candidate for a semi-automatic generation is the *business transaction*. The structure of a *business transaction* follows always the same pattern. However, instances of *business transactions* differ from each other with respect to the exchanged information, the names of the activities and the participating roles. The *business transaction announce waste transport* has been generated by the UMM Add-In using the form depicted in figure 4. Since every *business transaction* has a corresponding *business transaction use case*, the UMM Add-In offers the modeler a list of all *business transaction use cases* already defined in the BRV. Once a use case is selected, the modeler is presented with possible *participating roles* and *information envelopes*. Thereby the modeler saves a lot of time instead of manually creating each *business transaction* manually from scratch.

3.4 Validation of a UMM model

With the increasing complexity of a UMM model an inexperienced user often loses the overview and possible flaws

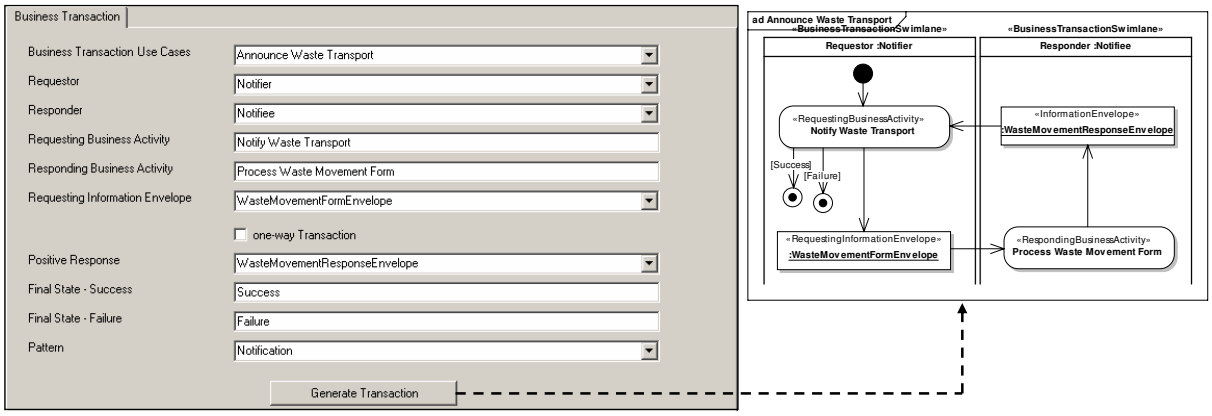


Figure 4: Semi-automatic generation of a business transaction

in the model remain undetected. A foible in the model such as a missing connector might not be an issue if the model is just used to communicate the structure of the inter-organizational business process between humans. If however, the model is used to derive business artifacts in an automated manner, a syntactically valid model is a prerequisite. As shown in the next chapter the UMM model can for instance be used to generate BPEL or BPSS documents. If the model is not correct the generator will most likely produce invalid documents or the generation fails completely.

Given these prerequisites, the UMM Add-In has a validation feature that allows checking the formal validity of a given UMM model. The validator is based on the OCL constraints incorporated in the UMM foundation module [25]. It however, does not interpret the OCL constraints directly but the logic is incorporated in the program code. The validator feature of the UMM Add-In allows two types of validation runs - bottom-up validation and top-down validation. The former one is used to validate sub-packages of a given UMM model and the latter one is used to perform a validation run on the whole model. After a validation run the errors found in the model are presented to the modeler. By double-clicking on an error message a detailed description of the error is shown.

Apart from the possibility to validate a given UMM model the validator will also help new users to get acquainted to the UMM standard. Many users do not read the whole specification but start modeling from scratch on a try and error basis. Especially the bottom up validator will gradually help the inexperienced user to create a valid UMM model and antagonize the argument, that UMM is too hard to learn or apply.

3.5 Generating process specifications for B2B information systems

In our UMM Add-In, we support the derivation of process descriptions for Web Services and ebXML - the two major approaches for implementing SOAs. In case of Web Services, we support the Business Process Execution Language (BPEL), for ebXML the tool generates choreographies according to the Business Process Specification Schema (BPSS). On a conceptual level, mappings from UMM to BPSS [8] and from UMM to BPEL [7] [9] have already been proposed. In our tool, the implemented transformation algorithms follow these approaches. The generated process specifications represent the local orchestrations which are derived from the global UMM choreography. In a real world scenario these process specifications are taken and fed into the business service interfaces of the participating business partners. Because each specification was derived from the global UMM choreography, the two business service interfaces participating in the inter-organizational business process are complementary and electronic business can be conducted in an automated manner.

3.5.1 Generating BPEL from UMM

BPEL describes the flow of a business process as a sequence of interactions between Web Services. An interaction is represented by an activity pointing to the respective service. Business partners participating in a process might provide services to the process and consume other partner's services. BPEL denotes the sequence of such Web Service calls, maps service calls to concrete Web Services and collates service consuming and service providing to the process participants. Thereby, BPEL describes the respective business process from a specific partner's point of view. Considering a business process internal to a company, whereby the company fully controls the logic of the process, the BPEL approach works fine. The process can easily be composed by orchestrating the required services in order to achieve the desired output. However, considering a complex B2B scenario that requires the participants to agree on process choreography, an approach whereby each participant describes its own view on the process in isolation will fail. In other words, if each business partner of a collaborative process describes its own BPEL the resulting process descriptions characterizing the same process will most likely not match. Thus, having BPEL ap-

plied in a B2B environment necessitates first a global view on the process, like UMM provides it, in order to gain complementary process descriptions.

In order to support this approach, the UMM Add-In allows the generation of partner-specific BPEL processes from a global UMM choreography. Since services are referenced by their WSDL in BPEL, the tool generates also a WSDL for each partner, describing the set of services he or she has to provide. Due to space limitations, we do not show the actual BPEL result.

3.5.2 Generating BPSS from UMM

BPSS denotes the choreography of a B2B business process for ebXML environments. It captures the flow of a collaborative process from a global viewpoint describing how each participant has to act in order to fulfill his or her part of the business process. BPSS is defined as an executable subset of the UMM specifying a B2B process by the concepts of a business collaboration consisting of business transactions. Although the BPSS specification does not mandate using UMM to gain BPSS process definitions, its use is recommended. Due to its alignment to UMM, BPSS allows the configuration of a business service interface according to the rich semantics of UMM business collaboration models. Given the space limitations the BPSS output is not show either.

4 UMM in the WeGo project context

The WeGo project *Enhancing Western Balkan eGovernment Expertise* [3] targets to boost eGovernment awareness and knowledge in order for Western Balkan Countries (WBC) to reach higher productivity and EU conformance. The major aim of the project is the transfer and the successful adoption of eGovernment best practices and knowledge. The project itself is split up into four distinctive yet complementary pillars namely: We-Go Interoperability Framework, We-Go Application Trials, We-Go Academies and We-Go Knowledge Net.

One of the main objectives of the project is the establishment of the WeGo Interoperability Framework with focus on transactional cross-border services and EU alignment. A second major task is the transfer of best practices for corresponding transactional application domains using demonstration prototypes. Services addressed by the project include eCustoms - New Computerized Transit System (NCTS), eJustice cross-border cases - Automation of Court Procedures (ACP), European Companies Register (ECR) and European Land Information System (EULIS), eAdministration - Electronic Filing System (implementation of paperless government) and eTrade Facilitation for European Waste Transport (EUDIN). Another goal is the establishment of eGovernment Academies to support and complement the Interoperability Framework and demonstration prototypes efforts and to establish a solid foundation for future regional eGovernment course program deployments. Finally the project aims on the establishment of an eGovernment Knowledge Net built by a system of federated registries to give easy access to relevant pieces of information, such as specifications or complete services including the solution and documentation. The WeGo Knowledge Net will encourage the re-use of existing approaches that will increase the likelihood of interoperability out of the box.

The analysis phase for the WeGo Interoperability Framework showed that there is a high need for clear and structured modeling and description of processes in the eGovernment area. Governments face interoperability issues when establishing cross-organizational services between different public authority bodies within the same or even different countries or services including European Union institutions. As mentioned above, cross-national services are covered by the WeGo project, e.g. in the eJustice and eCustoms area. The success of eGovernment heavily depends on the appropriate design and specification of processes in the application domain. Although the requirements in the various eGovernment domains may vary (as clearly indicated by the WeGo project) the specification of eGovernment processes should follow general, well-accepted design principles, such as UMM. A uniform structure and modeling of inter-organizational processes helps to increase business process interoperability. UMM is qualified to support that effort since it is a methodology for unambiguous definition of inter-organizational business processes. As an example, the eTrade Facilitation for European Waste Transport (EUDIN) system is not only based on IT-platform standards such as XML and Web Services, but also on standards describing the business logic of the system. One of the latter is UMM that has been used to define the business processes.

UMM is powerful but may be a challenge for users not familiar with it. Knowledge required to make use of UMM is made available through the WeGo Knowledge Net by providing specifications, tutorials or complete business process models for re-use. Another channel for the distribution of UMM know-how are the WeGo Academies.

5 Related Work

Over the last couple of years, a lot of methodologies for modeling business processes have been developed. Some of them are based on special notations often defined by standardization bodies. Others customize the UML for business process modeling needs. Traditionally, business process modeling focuses on modeling business processes internal to an organization fulfilling customer needs. More recent approaches also take inter-organizational business processes into account. Another criteria for distinguishing business process modeling approaches is its binding to the supporting IT infrastructure. Some approaches are mainly used in the requirements specification phase to support the communication with business domain experts. Resulting

models usually hide implementation complexity and are on a rather abstract level. Other approaches are more implementation oriented and rather provide a graphical interface for workflow languages or Web Service orchestrations/-choreographies.

Recently, the Business Process Modeling Notation (BPMN) [18] has attracted a lot of attention. BPMN is standardized by the Object Management Group (OMG) in order to enhance a uniform modeling notation for business processes, which is understandable by all business users - from the business analysts to the technical developers. In order to realize this goal, BPMN incorporates aspects of already advanced modeling notations (e.g. UML activity diagrams, IDEF [15], ebXML BPSS [27], RosettaNet [20], etc.). In order to close the gap between the business process design and the business process implementation, BPMN has a standardized mapping from BPMN to BPEL (Business Process Execution Language) [5].

Another very popular notation in business engineering are the Event-driven Process Chains (EPCs). EPC is a business process modeling language, focusing on control flow dependencies of activities in a business process. It is utilized in the ARchitecture of Integrated Information Systems (ARIS) by Scheer [21] as the central method for the conceptual integration of the functional, organizational, data, and output perspective in information systems design.

In addition to special business modeling notations, some approaches extend the UML meta model for this purpose. Initially, UML was designed to capture requirements for object oriented software systems in a graphical manner. Nevertheless some UML concepts, like UML activity diagrams, may be adopted for business process modeling. Korherr and List propose a UML profile for modeling business processes considering business process goals and performance measures [12]. Since the UML 2 activity diagram does not include such concepts, the UML meta model has been extended by a set of stereotypes and tagged values. On the one hand side this approach focuses on integrating measurable values into a business model at the conceptual layer. On the other hand side it specifies a mapping of UML models to BPEL, in order to make these values available for execution and monitoring. However, this approach is developed for modeling business processes internal to an organization. Further UML based approaches restricted to an internal view are proposed in [14] and [19].

For the purpose of representing and managing B2B business processes considering an inter-organizational view, Kim proposes a UML 1.x based modeling approach [11]. He uses activity diagrams for modeling collaborative processes as a flow of transactions in order to create an ebXML compliant business process specification. A transaction is a message exchange between two business partners and is represented as an activity in the activity graph. The flow of data exchanged between business partners is then captured in sequence diagrams.

A more IT-oriented approach is proposed by Tyndale-Biscoe et al. [22]. The authors present a UML 1.4 profile for the integration of business processes and software development used in an EU funded project. The UML profile provides the mapping between real world business concepts and software artifacts.

Kramler et al. [13] use UML 2 for depicting the choreography of Web Services. Unlike traditional business process modeling, additional requirements must be considered for service collaborations - e.g. security management or transaction management. In their paper the authors split the models into a layered architecture - collaboration, transaction, and interaction level, in order to compare these levels of granularity with the eCo framework [6]. The modeling technique is based on the considerations for a mapping to the Business Process Specifications Scheme (BPSS) and the Business Process Execution Language (BPEL).

6 Conclusion and Outlook

In this paper we have presented UN/CEFACT's Modeling Methodology (UMM), a language for designing B2B processes. UMM is defined as a profile for UML 1.4. In other words, the UML meta model has been extended by a set of stereotypes, tagged values and constraints in order to adjust UML for the semantics required by B2B. In theory, UMM models can be created by any UML tool. However, as we outlined there is a significant need for a customized UMM tool considering the UMM-specifics. Therefore, we developed the UMM Add-In - a plug-in for the commercial UML tool Enterprise Architect. This tool supports business analysts in creating UMM compliant business collaboration models. The functionality of the Add-In was characterized in the main part of this paper.

UN/CEFACT's Modeling Methodology in combination with our UMM Add-In is already in use in a couple of - mostly eGovernment related - projects for capturing collaborative business processes. The German Government uses UMM in order to define business collaborations between the local and federal governments. In this project a so-called XÖV Framework [4] is responsible for a harmonization of different XML based standards used in the public administration by using the concepts of UMM and Core Components. A further project using UMM has been set up by the Australian Government and is called GovDex [1]. The Department of Finance and Administration uses UMM and Core Components in order to improve the efficiency and effectiveness of business process collaborations across policy portfolios (e.g., taxation, human services, etc.) and administrative jurisdictions. The Canadian Administration is trying to apply the principles of UMM to describe the activities of governmental agencies. This work is known as the Governments of Canada Strategic Reference Model (GSRM) [2].

Future research efforts will concentrate on the development of the successor of the current UMM standard - UMM 2.0. At the time when UMM 1.0 was developed the support for UML 2.0 in modeling tools was poor. However today

UML 2.0 is state of the art and therefore UMM 2.0 will be based on UML 2.0. Current research is also conducted in the field of business document modeling. The current version of the UMM Add-In supports the Core Component Technical Specification 2.01. The next version of CCTS (3.0) is currently under development. It is planned that future versions of the UMM Add-In support the latest CCTS standard as well.

During the We-Go project UMM will be used to model eGovernment processes in the Western Balkan Region. Valuable feedback from this usage will be used to further improve the UMM and its tool support. Lessons learned from finished UMM projects will be analyzed and suggested changes will be introduced in UMM 2.0.

References

[1] GovDex Framework. http://www.govdex.gov.au.

[2] Governments of Canada Strategic Reference Model. http://www.tbs-sct.gc.ca/index_e.asp.

[3] We-Go EU project. http://www.wego-project.eu, Proposal No 045472, Contract No 045472.

[4] XÖV Framework. http://www.osci.de (german).

[5] BEA, IBM, Microsoft, SAP AG and Siebel Systems. *Business Process Execution Language for Web Services*, May 2003. Version 1.1.

[6] eCo Working Group. *eCo Architecture for Electronic Commerce Interoperability*, 1999.

[7] B. Hofreiter and C. Huemer. Transforming UMM Business Collaboration Models to BPEL. In *Proceedings of OTM Workshops 2004*, volume 3292, pages 507–519. Springer LNCS, 2004.

[8] B. Hofreiter, C. Huemer, and J.-H. Kim. Choreography of ebXML business collaborations. *Information Systems and e-Business Management (ISeB)*, June 2006.

[9] B. Hofreiter, C. Huemer, P. Liegl, R. Schuster, and M. Zapletal. Deriving executable BPEL from UMM Business Transactions. In *Proceedings of the 2007 International Conference on Services Computing*, pages 178–186. IEEE, 2007.

[10] ISO. *Open-edi Reference Model*, 2004. ISO/IEC JTC 1/SC30 ISO Standard 14662, Second Edition.

[11] H. Kim. Conceptual Modeling and Specification Generation for B2B Business Processes based on ebXML. In *SIGMOD Rec., vol. 31, no. 1, pp. 37-42*, 2002.

[12] B. Korherr and B. List. Extending the UML 2 Activity Diagram with Business Process Goals and Performance Measures and the Mapping to BPEL. In *Proceedings of the ER-Conference on Conceptual Modeling (Workshop BP-UML'06)*. Springer, Nov. 2006.

[13] G. Kramler, E. Kapsammer, G. Kappel, and W. Retschitzegger. Towards Using UML 2 for Modelling Web Service Collaboration Protocols. In *Proceedings of the First International Conference on Interoperability of Enterprise Software and Applications (INTEROP-ESA'05)*, Feb. 2005.

[14] B. List and B. Korherr. A UML 2 Profile for Business Process Modelling. In *ER 2005 Workshop Proceedings*, 2005.

[15] R. J. Mayer, P. C. Benjamin, B. E. Caraway, and M. K. Painter. A Framework and a Suite of Methods for Business Process Reengineering. *Business Process Reengineering: A Managerial Perspective*, pages 245–290, 1995.

[16] OASIS, UN/CEFACT. *ebXML - Technical Architecture Specification*, February 2001. Version 1.4.

[17] Object Management Group (OMG). *Unified Modeling Language Specification*, Apr. 2005. Version 1.4.2, http://www.omg.org/docs/formal/05-04-01.pdf.

[18] OMG - Object Management Group. *Business Process Modeling Notation Specification 1.0*, 2006.

[19] M. Penker and H.-E. Eriksson. *Business Modeling With UML: Business Patterns at Work*. Wiley, 2000.

[20] RosettaNet. *RosettaNet Implementation Framework: Core Specification*, December 2002. V02.00.01.

[21] A.-W. Scheer. *ARIS - Business Process Modeling*. Springer, 2000.

[22] S. Tyndale-Biscoe, O. Sims, B. Wood, and C. Sluman. Business Modelling for Component Systems with UML. In *Proceedings of the Sixth International Enterprise Distributed Object Computing Conference (EDOC'02*. IEEE Press 2002, 2002.

[23] UN/CEFACT. *Electronic Data Interchange*, September 2006. D.06A.

[24] UN/CEFACT International Trade and Business Process Group (TBG14). *UN/CEFACT Common Business Process Catalog*, Nov. 2003. Version 0.95.

[25] UN/CEFACT Techniques and Methodologies Group. *UN/CEFACT's Modeling Methodology (UMM), UMM Meta Model - Foundation Module*, Oct. 2006. Technical Specification, http://www.unece.org/cefact/umm/UMM_Foundation_Module.pdf.

[26] UN/CEFACT TMG. *Core Components Technical Specification - Part 8 of the ebXML Framework*, November 2003. v2.01.

[27] UN/CEFACT TMG. *UN/CEFACT - ebXML Business Process Specification Schema*, 2003.

[28] UN/CEFACT TMG. *BCSS - UML Profile for Core Components based on CCTS 2.01*, October 2006. Candidate for version 1.0.

Games for Learning and Learning from Games

Maja Pivec
Information Design, FH JOANNEUM, University of Applied Sciences, Graz, Austria
E-mail: Maja.Pivec@fh-joanneum.at, http://mpivec.radical.ac.nz

Paul Kearney
Deakin University, Australia/UNITEC School of Information Technology,
Auckland, New Zealand
E-mail: Paul@radical.ac.nz, http://pkearney.radical.ac.nz

Keywords: game-based learning, recursive loops of learning, games for learning

Received: June 12, 2007

Abstract: This paper details a model of game-based learning and suggests how this can be applied to both the playing of computer games and learning within the classroom environment. The authors document the results from a University level course, created in the form of a role-play for designing educational games, and highlight the student's attitudes and beliefs towards game design as a career. They also suggest that educational games can be used successfully for the transference of knowledge to domains outside the world of computer games, and highlights several case studies in the area of health and medicine.

Povzetek: V prispevku je predstavljen model učenja na osnovi iger in možnosti uporabe modela v različnih okoljih: pri igranju iger kakor tudi pri uporabi iger za učenje v razredu.

1 Background and introduction to game-based learning.

Over last few years an emerging trend of games in the area of e-learning has been observed. From early isolated reports on conferences and books reflecting about possible application of digital games for learning (Gee, 2003), more and more practitioners and researchers embraced the idea, including the e-learning community. In 2006 one of the biggest European e-learning conferences, Online Educa in Berlin, introduced a special game track. The two-day session hosted an open discussion between academics, teachers and industry practitioners, focusing on the potential of game-based learning in Universities and lifelong learning institutions and possible software solutions.

The discussions are primarily focused on Pros and Cons of the application of games for learning, trying to find answers to *Why don't we use games more often in classrooms?* Often it is pointed at the difficulty to find games that cover the curricular topics, the low tolerance of the environment towards the games where the games are often perceived as unserious activity, with some lecturers fearing that the learning objectives wouldn't be reached, and others might encounter difficulties with technical resources that schools don't have. Another important factor is the quality aspect of the games for learning where games should have an explicit learning purpose and can be used, adapted and adopted for supporting, improving and fostering learning processes (SIG-GLUE).

Kasvi (2000) lists the seven requirements suggested by Norman (1993) for an effective learning environment to be to:

1. Provide a high intensity of interaction and feedback;
2. Have specific goals and established procedures;
3. Be motivational;
4. Provide a continual feeling of challenge, not too difficult to be neither frustrating nor too easy to create boredom;
5. Provide a sense of direct engagement on the task involved;
6. Provide the appropriate tools that fit the task; and
7. Avoid distractions and disruptions that destroy the subjective experience.

Kasvi (2000) suggests that computer games fulfill all of these requirements and believes that they "satisfy them better than most other learning mediums" (p.6). However, it is very difficult to find a game that includes a learning curriculum that is appropriate for different schooling levels. Popular games like 'Maths Blaster' from Vivendi Universal, has captivated children but only targets ages 8 to 9 years. Even if the game were upgraded to include a higher level of mathematics, it would be doubtful that today's 14-year-old student would play this type of game. But take a constructivists point of view and ask the same student to design an educational game, and the response would be quite different, as described in section 3.

Today's students are captivated by computer and console video games. Humans have always used games of all types for learning - from playing with blocks for counting skills, to flight simulators for skills of a more specialised nature (Pivec, 2006). Although the skills involved when playing games differ dramatically from

those needed to create one, players exhibit the same addictive nature seen in a person that is driven to succeed. A computer game can take anywhere between 3 months and 3 years to create. From the initial concept, the design, the coding, testing and error correction, through to the artwork, the music, packaging, promotion, and distribution, developers must stay focused and committed to the project throughout this entire time, often doing tedious tasks but always learning new and innovative techniques to do their craft. These people are usually young adults and have been avid game players themselves. They learn in a different way from earlier generations and are often motivated by instant feedback and reward of success.

Game-based learning can be applied as additional option to classroom lecturing. The intention of game-based learning is to address new ways of ICT based instructional design and at the same time to provide learners the possibility to acquire skills and competencies later required in the business world. By means of digital games and especially of digital educational games learners should be able to apply factual knowledge, learn on demand, gain experiences in the virtual world that can later shape their behavioral patterns and directly influence their reflection, etc.

2 Recursive loops of Game-Based Learning.

Let us consider, based on the example of an educational adventure game, how and when learning occurs when learners interact e.g. play a game. The main characteristic of an educational game is the fact that instructional content is blurred with game characteristics. The game should be motivating, so that the learner repeats cycles within a game context. While repeating e.g. playing a game, the learner is expected to elicit desirable behaviors based on emotional or cognitive reactions which result from interaction with and feedback from game play.

The purpose of an adventure game is entertainment or edutainment. In adventure games there are very complex environments i.e. microworlds, with no deterministic problem representation. An examples of typical edutainment game is Chemicus (by Heureka-Klett publisher; or TIVOLA for the US market), a puzzle-adventure game for self directed learning of chemistry. Similar to Chemicus one can find an entire series of titles e.g. Physicus, Hystorion, Informaticus, etc. by the same publishers built upon the same game concept.

Adventure games use intrinsic motivation of the player to explore the game world. Intrinsically motivating games incorporate learning activities in this game world. To increase immersion of the player, the game offers an extensive story at the beginning, often related to some murder or mystery. Game characters have to solve the mystery by solving a number of interrelated problems. In each case the problems are part of the game and players are motivated to seek for knowledge to provide a solution in order to continue with the game. In the described game, enjoyment is strongly related to the learning activity, which can be viewed as a desirable outcome.

Commercial computer games are known for creating social environments and cult followings surrounding the gameplay, the character attributes, and player's abilities, and this is where affective learning occurs (Kearney and Pivec, 2007). Garris et al., (2002) describes affective learning as including "feelings of confidence, self-efficacy, attitudes, preferences, and dispositions" (p.457). The skill-based learning appears to comfortably fit within the micro game cycle (figure 1), or levels within the game. For example, Rosser et al., (2007) found that the playing of commercial action games improved the surgical skills of laparoscopic physicians and decreased their error rate. There was no documented debriefing session for Rosser's study and it is assumed that the development of technical or motor skills occurs within the game itself.

Figure 1 also shows how player ability and experience affects the challenge element and the level of learning (Zone of Proximal Development), and how the level of cognitive challenge can be appropriate for the learner's current abilities. The model shows the inclusion of instructional design and game characteristics as critical elements of a game to enable the achievement of the learning outcomes, as well as the additional factor of player abilities. Defining learning as the acquisition of knowledge or skills, suggests that Game-Based Learning is the vehicle that fosters the acquisition of the learning outcomes. The model includes a time element to allow the player to progress through the game increasing their knowledge and acquiring new levels of ability. This suggests that knowledge, declarative, procedural, and strategic is acquired over time and abilities or skills are incremented through experience.

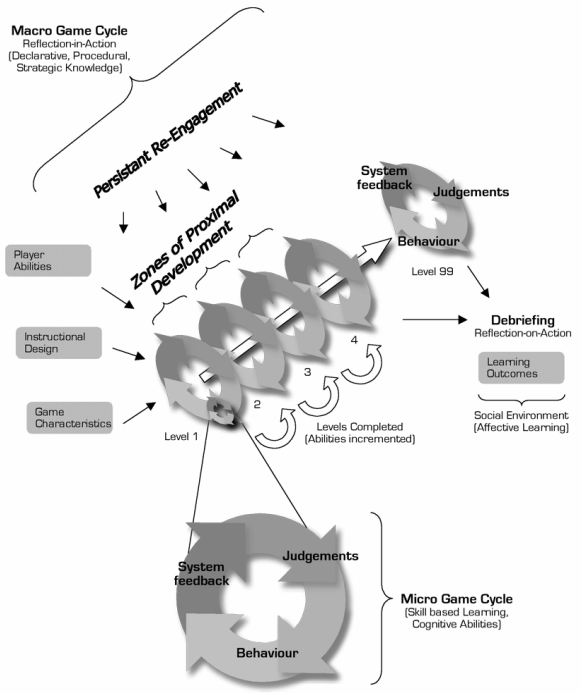


Figure 1: Recursive loops of Game-Based Learning (Kearney & Pivec, 2007).

This model can also be applied to role-play within the classroom. As the student's abilities are added to, through tuition or guided instruction, their knowledge and skill level is incremented and they move to the next level, or next phase of the project. The role-play course on game design detailed in the next section, was scaffolded in such a way that the students added to their game design concept as their knowledge and skill increased.

3 Game about the game design: role-play in classroom.

This section documents an educational game design course created by authors and taught to 75 information design students at the University of Applied Sciences Joanneum in Austria, where we wanted to introduce this topic to the coming generation of potential game designers and make them aware of this new discipline along with its specifics. The challenge for students was to create a concept proposal for a publisher of educational games. Based on the course work and results, we analysed how students perceived the area of educational games for the use of teaching and as a career path.

The class was a role-play itself i.e. game about designing a game, where students had to work in teams, create a game design company and take a specific role and responsibilities within the team e.g. game producer, game developer, programmer, etc. to contribute to the task accomplishment. The progress of the work along with the problems they encounted were documented within the company blogs (see as an example http://legalaliengames.blogspot.com/, blog of "the best in show" group)

The course covered topics including the process of commercial game design, taking into consideration the pedagogical design required to achieve the desired learning outcomes. When we design games for learning both the target audience and the learning outcomes have to be considered at the initial conception of the game. In this way teachers can easily recognise the value of this resource and the possibilities to include such games in the curriculum. Aspects of educational game design are tackled more in detail in (Pivec, Koubek & Dondi, 2004).

Figure 2: Student Designs

The game concepts differ in excellence from the innovative use of technology to their possible market potential. The class finished with the presentation with the Golden Pineapple awarded concepts (Golden Pineapple Award, 2006). Two of the awarded concepts were focused on the medical content (Figure 2 & 3). Anaphylactic from Dudary Entertainment is a real time strategy game introducing the principles of immune system of human body. Keep Me Alive from Stardust Enterprises is an ICQ game focusing on various infectious diseases and how to prevent and treat them. It also has the potential to include real pharmaceutical products as well as relevant medical advice.

Students were surveyed both, before and after the completing the course on their opinion on games in general and regarding the potential of application of games for learning. We also inquired the motivational momentum of designing a game in terms if they were more motivated and achieved better learning results. Based on this survey we also wanted to assess if they saw educational game development as a possible career path.

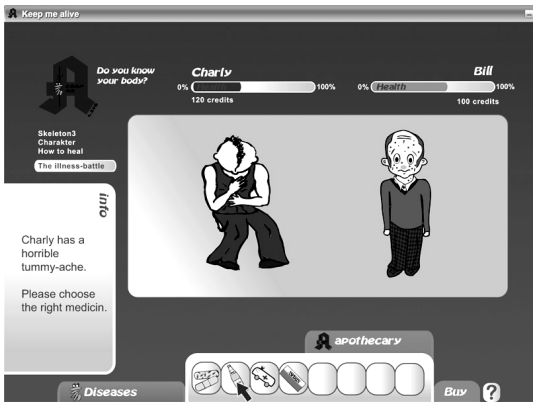


Figure 3: Student Designs

On the post survey, 66% of the students agreed that designing educational games was a highly motivational topic and suggested that they now felt competent enough to write a professional educational game concept document. They also agreed the designing educational gems could provide future career opportunities, however only 35% of them would consider this for their own career. The majority of the students found the course to be successful with 70% of the students enjoying the topic despite not considering themselves to be game players. Those who did play computer games, only did so for recreation and had not involved games with any of their schooling. However, upon completion of the course, 60% of the students suggested a preference for using games to learn.

4 The Application of Game-Based Learning.

With the intention to outline the potentials of application of games in the area of medicine (as a serious discipline in contrast to the computer games, that are often seen only as a leisure activity or even as a waste of time),

some known and documented cases of application of game-based learning targeting various user groups are presented. The cases vary from an educational game created for interdisciplinary learning, to context based environments supporting the application of isolated knowledge for medical and veterinarian students, and the application of commercial-off-the-shelf games (cots) to improve the laparoscopic performance, embedded in the curricula.

Suzanne de Castell and Jennifer Jenson from Canada created Contagion, a role-playing adventure game fostering interdisciplinary learning and targeted at children aged 10 – 15 (de Castell and Jenson, 2006). The game is based on traditional school subjects and related subject fields like technology, biology and medical sciences, as well as human and social sciences. The goal of the game is twofold. On one hand, the game should introduce health related topics and educate players by means of "serious play" about diseases, such as Severe Acute Respiratory Syndrome (SARS), West Nile Virus (WNV), Avian Flu, and Acquired Immune Deficiency Syndrome (AIDS), and possible preventive behaviours. On the other hand, the game also provides a career preparation environment; where players can learn about and role-play various occupation of interest e.g. community health officer, physician, or a medical researcher. The player entering the game world chooses one of these roles that effect the development of the game play and the point of the view on the situation throughout the game. In the game the player is confronted with the situation of medical and humanitarian crisis, and acts out the situation differently based on the respective role. The majority of the learning is based on active exploration.

At the University of Edinburgh, students interact with virtual patients from their first year of study until completion. The virtual patients are related to various curricular topics blurred with narrative elements thus creating a realistic context (Begg et al., 2006). Each student interacts with the same virtual patients, e.g. George, for several times throughout their study. His condition gets more complicated as they progress in their studies. The aim of George is to provide an opportunity to apply concepts learned in isolation e.g. social and cultural factors of health and communication skills. By interacting with these virtual patients students are role playing "to be a doctor", until the end of their education when they become doctors. Labyrinth is similar application based on the virtual patients and realistic scenarios that has been created for the College of Medicine and Veterinary Medicine's Learning Technology Section at the University of Edinburgh. The scenarios are focused on decision-making i.e. the students decisions and courses of action influence further development of the scenario. At the beginning the student is placed in the role of being in charge of an admissions unit at the start of the night shift. The student is confronted with the situation based on a short descriptive text and asked what to do next. They are offered a set of choices; some of them are more appropriate than others. Based on the development of the scenario they get feedback on their reflection and choices made. Technology i.e. virtual scenarios, have the advantage of being able to restart the session and try out the "what if" reflections repeatedly.

Newly published research suggests that video games may be a teaching tool for training of laparoscopic skills (Rosser et al., 2007). The study involved thirty three male and female surgeons of various specialities and was centred at Rosser Top Gun Laparoscopic Skills and Suturing Program where the goal is to build skill sets that enable surgeons to function effectively in the video-endoscopic surgical environment. One part of the study included playing three cots video games. At the end of the study the results of laparoscopic performance were grouped in the categories based on the gaming experience i.e. past players, current players and demonstrated skills in the games a part of the study and then compared to the laparoscopic results of non-players. The published results showed that current video game players made 32% fewer errors (P=0.04), performed 24% faster (P=0.04), and scored 26% better overall (time and errors) (P=0.005) than their non-paying colleagues (Rosser *et al.,* 2007). Based on the carried out research Rosser argues that video games "may help thin the technical interface between surgeons and screen-mediated applications" thus contributing to better performance in laparoscopic surgery in terms of faster completion and fewer errors.

5 Conclusions.

In many cases, the application of serious games and simulations for learning provides an opportunity for learners to apply acquired knowledge and to experiment, get feedback in form of consequences thus getting the experiences in the "safe virtual world". There are specific educational domains where game-based learning concepts and approaches have a high learning value. These domains are interdisciplinary topics where skills such as critical thinking, group communication, debate and decision making are of high importance. Such subjects, if learned in isolation, often cannot be applied in real world contexts.

Games can provide the motivation to learn, increasing the likelihood that the desired learning outcomes will be achieved. Learning is defined as the acquisition of knowledge or skills through experience or practice, and what better way to learn than through a game.

Acknowledgments.

Many thanks go to the students of FH JOANNEUM, Department of Information Design, IND04 and IND05 for their willingness in participation at the pre and post survey as well as for their work during the courses.

References.

[1] Begg M., Ellaway R., Dewhurst D. & Macleod H. (2006). Virtual Patients: considerations of narrative

and game play. In Burmester M., Gerhard D. & Thissen F. (Eds.).*Digital Game Based Learning, Proc. of the 4th International Symp. For Information Design*, Stuttgart Media University., pp. 101 - 113

[2] de Castell, S. & Jenson, J. (2006). How Content Matters: Rethinking Educational Games. In Pearson E., & Bohman P. (Eds.). *Proc. of ED-MEDIA'06*, Orlando, Florida, USA, 2006, pp. 1802-1808

[3] Garris, R., Ahlers, R. & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, **33**(4), 441-467.

[4] Gee J.P. (2003). *What video games have to teach us about learning and literacy*. New York, Palgrave Macmillan.

[5] Golden Pineapple Award (2006). Lehrgang Game-based Learning: Abschlusspräsentationen. Available from 06. 04. 2007 from http://informations-design.fh-joanneum.at/log-logbucheintrag.php?id=192 (06. 04. 2007)

[6] Kasvi J. (2000) *Not Just Fun and Games - Internet Games as a Training Medium.* Cosiga - Learning with Computerised Simulation Games. pp.23-34. Available from http://www.knowledge.hut.fi/people/jkasvi/NJFAG.PDF (27 June 2003)

[7] Kearney, P. & Pivec, M. (2007). Recursive loops of game based learning. In Montgomerie C., & Seale J. (Eds.). *Proc. of ED-MEDIA'07*, Vancouver BC, Canada, 2007, pp. 2546 – 2553.

[8] Norman, D. (1993). *Things that make us smarter: Defending Human attributes in the age of the machine.* New York, Addison -Wesley

[9] Pivec M. (Ed.) (2006). *Affective and emotional aspects of human-computer interaction; Game-Based and Innovative Learning Approaches. Vol.1: The Future of Learning* (IOS Press, 2006), ISBN 1-58603-572-x

[10] Pivec, M., Koubek, A. & Dondi C. (Eds.) (2004). *Guidelines on Game-Based Learning*. Pabst Vrlg., ISBN: 3899671937

[11] Rosser J.C., Lynch P.J., Cuddihy L. & Gentile D.A. (2007) The Impact of Video Games on Training Surgeons in the 21st Century. *Arch Surg.* **142** (Feb. 2007);181-186 (www.archsurg.com).

[12] SIG-GLUE: Special Interest Group for Game-based Learning in Universities and Lifelong Learning; project web-page. Available from http://www.sig-glue.net (06. 04. 2007)

Prediction of Missing Data for Ozone Concentrations Using Support Vector Machines and Radial Basis Neural Networks

Biljana Mileva-Boshkoska and Mile Stankovski
Department of Automatics, Faculty for Electrotechnics and Information Technologies,
Karposh 2, bb, 1000 Skopje, Republic of Macedonia
E-mail: biljanamb|milestk@feit.ukim.edu.mk

Keywords: air modelling, support vector machines, neural networks

Received: July 6, 2007

In this paper we present results from prediction of data for ozone (O_3) concentrations in ambient air by using the modelling techniques of support vector machines (SVM) and radial basis neural networks (RBF NN). The predictions are performed for two short periods of time: for 24 hours and for one week in August and in December in 2005, in Skopje, Macedonia. The built SVM models use different kinds of kernels: polynomial and Gaussian kernels and the best values of the free parameters of SVM kernels are chosen by examining a range of values for each of the free parameters. This is the first attempt in Macedonia for prediction of concentrations of any air parameters in the ambient air.

Povzetek: Podana je analiza ravni ozona v Makedoniji z metodami strojnega učenja.

1 Introduction

In the process of EU integration, Republic of Macedonia had to harmonize environmental legislation with European one. According to the new Macedonian legislation for air quality (Law on ambient air quality, Official Gazette of Republic of Macedonia, no 67/2004) the country is obliged to perform continuous monitoring of the ambient air throughout the whole territory of the country. For that reason, in Macedonia were installed fifteen automatic monitoring stations for gathering data for the air quality. However, mainly due to financial reasons, and technical problems in the maintenance of the monitoring stations, the data sets from the monitoring stations are not complete. According to the EU directives and Macedonian legislation, the country must fulfill 90% of the yearly measurements for the air quality on the measuring spots during one year. In order to fulfill the gaps in the data sets for air quality, we decided to use appropriate mathematical modeling technique, as a method that is allowed to be used by the EU directives.

In this paper we present the results obtained from filling in the existing gaps of the measured hourly data for the levels of ozone (O_3) in the ambient air for two short periods of time in the municipality of Karposh III, in Skopje, Republic of Macedonia. We process the two data sets for August and December, 2005 and we build statistical models for hourly predictions of concentrations for one day and for one week. Solution of the problem had to be generated in a simple manner and the used algorithm had to be applicable for similar problems e.g for prediction of concentrations of other air quality parameters.

One approach for prediction of hourly values is using neural networks for evaluating air parameters concentrations [1], [2], [3], [4]. SVM is another method that started in the late seventies [5], [6] and today is used for ambient air parameters prediction [7], [8], [9] and for time series forecasting [10] in the environmental applications.

For prediction of the O_3 levels, we use the modelling techniques based on SVM and Radial Basis Function (RBF) NN.

2 Overview of the whole process

Prediction of levels of O3 in ambient air is a complex process that consists of the following phases (Figure 1):

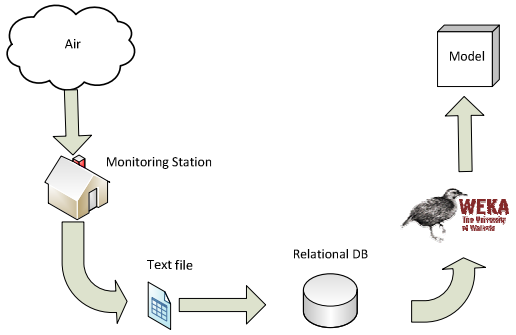


Figure 1 Overview of the whole process of prediction

- Measurement of the levels of parameters of the ambient air by automatic monitoring station.
- Transmission of the measured data via radio connection from the monitoring station to the textual data base situated in the Ministry for Environment and Physical Planning (MoEPP).
- Data processing and preparation of ARFF files that are recognized by the WEKA software.
- Electing tools (software) for modelling the data.
- Modelling using the software package WEKA.
- Comparison of the received models and choosing the one that gives the best prediction results.

3 Used techniques

3.1 Support Vector Regression

The concept of a maximum margin hyperplane only applies to classification. However, support vector machine algorithms have been developed for numeric prediction that share many of the properties encountered in the classification case: they produce a model that can usually be expressed in terms of a few support vectors and can be applied to non-linear problems using kernel functions.

Similar with linear regression, the basic idea here is to find a function that approximates the training points well by minimizing the prediction error. The crucial difference is that all deviations up to a user-specified parameter $\mathbf{x_i}$ are simply discarded. Also, when minimizing the error, the risk of overfitting is reduced by simultaneously trying to maximize the flatness of the function. Another difference is that what is minimized is normally the predictions' absolute error instead of the squared error used in linear regression. A user-specified parameter $\mathbf{x_i}$ defines a tube around the regression function in which errors are ignored.

SVM approximate the learning data set with a function given in a form of:

$$\mathbf{f(x)} = \sum_{\mathbf{i=1}}^{\mathbf{l}} \mathbf{w_i \phi_i(x)} + \mathbf{b} \tag{1}$$

meaning that the original data $\mathbf{x} \rightarrow \boldsymbol{\phi}(\mathbf{x})$ are mapped into high dimensional space and then construct an optimal hyperplane in this space. $\boldsymbol{\phi}(\mathbf{x})$ represents feature of the inputs, while $\mathbf{w_i}$ and b are coefficients. These are estimated by minimizing the risk function [10]:

$$\mathbf{R(f)} = \int \mathbf{c(x, y, f(x))dp(x, y)} \tag{2}$$

where c(x,y,f(x)) is cost function that determines how to penalize estimation errors based on the empirical data X [7]. Given that we do not know the probability measure dp(x,y) we can only use X for estimating a function f that minimizes R[f]. A possible approximation consists in replacing the integration by the empirical estimate to get so called empirical risk function

$$\mathbf{R_{emp}[f]} = \frac{\mathbf{1}}{\mathbf{l}} \sum_{\mathbf{i=1}}^{\mathbf{l}} \mathbf{c(x_i; y_i; f(x_i))} \tag{3}$$

A first attempt would be to find the function $\mathbf{f_0} = \mathbf{argmin_{f \in H} R_{emp}[H]}$) for some hypothesis class H. However if H is very rich, i.e. its capacity is very high as for instance when dealing with few data in very high dimensional spaces, this may be not such a good idea as it will lead to overfitting and thus bad generalization properties. Hence one should add a capacity control term, which in the SV case results to be $\left|\left|\mathbf{w}\right|\right|^{\mathbf{2}}$, which leads to regularized risk function

$$\mathbf{R_{reg}} = \mathbf{R_{emp}[f]} + \frac{\boldsymbol{\lambda}}{\mathbf{2}} \mathbf{\|w\|^2} \tag{4}$$

3.2 Kernels

A kernel is essentially a similarity function with certain mathematical properties, and it is possible to define kernel functions over all sorts of structures-for example, sets, strings, trees, and probability distributions.

The choice of kernel $\mathbf{K(x_i, x_j)}$ influences drastically on the performance of the SVMs depending on the problem considered. Several kernels are available for learning sand they have to satisfy the so-called Mercer's condition [9].

The most commonly used kernels are the Gaussian kernel

$$\mathbf{K(x_i, x_j)} = \mathbf{exp}^{-\left(\frac{\left\|\mathbf{x_i - x_j}\right\|^{\mathbf{2}}}{\mathbf{2\sigma^2}}\right)} \tag{5}$$

and the polynomial kernel

$$\mathbf{K(x_i, x_j)} = \mathbf{(x_i x_j + 1)^p} \tag{6}$$

which are also used for the purposes of this research.

4 Data processing

The data sets that are used are gathered by the national automatic monitoring network (AMN) by the MoEPP in Republic of Macedonia. As soon as the data are transferred to the central DB in MoEPP they are first validated, that is the missing and the unreal data are marked with -9999. We have picked a small period of time where we do not have missing data, that is the period 1-17 August and 1-17 December 2005. We used two different seasons because we wanted to show the difference of the predicted results from different models depending on the standard deviation of the input data.

The first phase is parsing of data and their storage in a relational data base. We convert the validated data into ARFF format that is recognized by the WEKA software that is used for the process of prediction of the O_3 levels.
In order to build models for prediction of O_3 levels, as input parameters we use the hourly data for the levels of NO_2, O_3, temperature and humidity for 10 days in a raw. The output function is following:

$$\boldsymbol{O_3(t) = f(NO_2(t-z) + O_3(t-z) + NO_2(t) + temp(t-z) + hum(t-z))} \quad (1)$$

We built eight different models for prediction of O_3 levels for $t - z$ hours, where $z = 1,2, \ldots, 8$.

For prediction of O_3 levels, first we build three types of models from which two are based on SVM, while the third one is based on RBF NN. In order to build the first two models, we use the existing functions in WEKA: SVMreg with polynomial kernel, where p=1 and SVMreg with RBF kernel, known as SVM with Gaussian kernel. For building the third model we use the function RBF with neural network which is also implemented in Weka. The three functions are used both for prediction of levels of O_3 for 24 hours and for one week. That way we get two groups of models. In the first group belong models for prediction of levels for 24 hours and in the second group belong models for prediction of levels for one week.
In order to decrease the total processing time for training the SVM we used the tool Explorer from WEKA that enabled us to distribute the whole process of learning of the model on three computers controlled by one "master" computer.

The results from the obtained models are compared. As a measure for deviation of the predicted results from the measured one we use the mean absolute error given by:

$$\boldsymbol{MAE = \frac{1}{n}\sum_{i=1}^{n} |\alpha_i - p_i|} \quad (2)$$

5 Results from modelling

When modelling with SVM, first we choose the best values of the free parameters of the kernels: C (factor of penalty, Figure 2), ε (Figure 3) and γ,which is connected to the speed parameter σ with the relation of $\gamma = \sqrt{\left(\frac{1}{\sigma^2}\right)}$ (Figure 4). To choose the values of the free parameters is the main difficulty when modelling with SVM. Taking into consideration that there are no general rules on determination of the values of the free parameters, it is necessary to determine the influence of the chosen value of the free parameter on the resulting error on the predicted results from the model. In this paper we use MAE for assessment of the deviation between the original measured data and the predicted data. In general, the smaller MAE, the better results the built models achieve.

Figure 2 presents the variations of MAE from the parameter C. The graph shows that the parameter C has very small influence on MAE and it is sensitive only on very small values for C, for example when $C \leq 0.001$. When increasing the values of C, the value of MAE steeply decreases until C recives values $C \geq 0.5$ when again parameter C makes very small influence on MAE. In general, in order one to guarantee a stable learning process, the value of the parameter C has to receive large values, for example C=100, as it is the case in this paper.

Figure 3 presents the variations of MAE from the values of the parameter ε. Parameter ε, like parameter C has small influence on the performances of the model for prediction of the ozone concentrations. The values of MAE are almost constant for values of the parameter $\varepsilon < 10^{-2}$ and $\varepsilon > 0.5$. In the models in which we use SVM the value of ε is should be small. In this research, we set the value of ε to 0.1.

Theoretically, the value of the speed parameter σ influences a lot on the prediction performances of the model. Very small ($\sigma \to 0$) or very large ($\sigma \to \infty$) values of σ may lead to bad prediction results. If $\sigma \to 0$, all training data become support vectors. In that case, when an unknown data occur as input at the SVM model, the SVM model will not be able to provide good predication results. From the other side, if $\sigma \to \infty$, all training data will be considered as one point and the SVM model may produce same results for any new input data to the model. Therefore, these two extreme cases should be avoided. We should note that both $\sigma \to \infty$ and $\sigma \to 0$ represent two approximate processes. In real applications, if $\sigma \ll \|x_i - x_j\|$ and $\sigma \gg \|x_i - x_j\|$ the extreme cases mentioned above will occur. Figure 4 presents the variations of MAE from the values of the parameter σ. Results in the **Error! Reference source not found.** show that MAE is large, when σ is small (for example $\sigma = 0.001$), than it decreases with increasing of σ and it reaches minimum for values of σ around 1. Figure 4 shows that MAE fluctuates when γ is in the range of [0.9, 1.1]; then it increases with increasing of γ, and finally it has tendency to become constant after γ reaches values $\gamma \geq 100$. For that reason, in practical applications only parameter γ (or σ) of the Gaussian kernel function has to be determined, while the two parameters C and ε may be set in advance by experience. In this application we set the value of γ to 0.5.

Once the best values for the free parameters C, γ (or σ) and ε are determined, the final step is to produce the models for prediction of the missing data for O_3.

In this paper, we have calculated the best values for parameters C, ε and γ for z=3. We have used the same values later in order to predict results for $z \neq 3$ ($z = 1,..,8$) in which cases the free parameters are not optimal. Although in those cases we do not use the optimal values for the free parameters, still the results obtained from models built with SVM are better than those from the models built with RBF NN.

Figures 5 – 8 show the results from the modelling. Each figure show the distribution of the original data and the distribution of predicted data obtained from three different models built with polynomial and Gaussian kernel and with RBF NN.

In August 2005, the data of the levels of O_3 are very close to each other i.e. the standard deviation is very small. Therefore the three models give similar results for prediction of O_3 levels (Figure 5 and Figure 6).

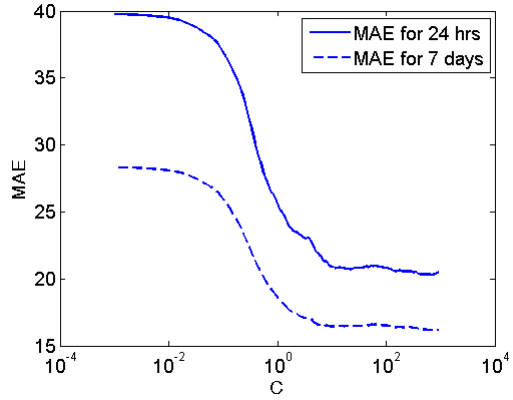


Figure 2: Variations of MAE from the parameter C for prediction of O_3 levels for 24h and for 7 days for August, 2005

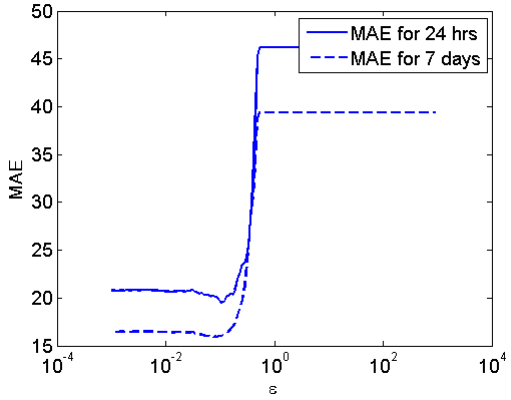


Figure 3: Variations of MAE from the parameter **ε** for prediction of O_3 levels for 24h and for 7 days for August, 2005

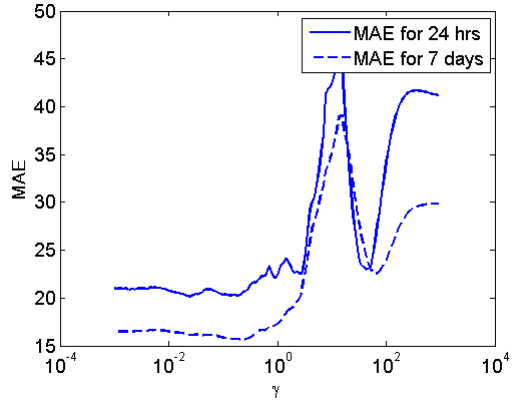


Figure 4: Variations of MAE from the parameter **γ** for prediction of O_3 levels for 24h and for 7 days for August, 2005

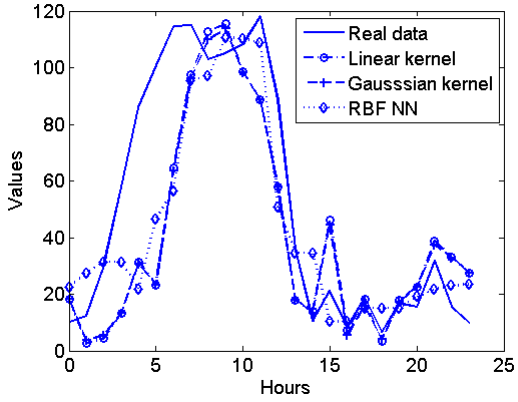


Figure 5: Prediction of levels of O_3 for 24 hours, for August 2005, for z=3

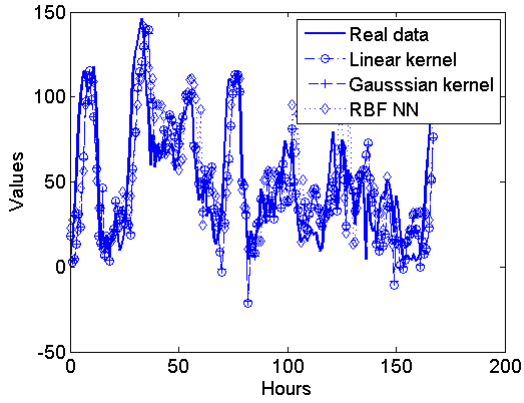


Figure 6: Prediction of levels of O_3 for one week, for August 2005, for z=3

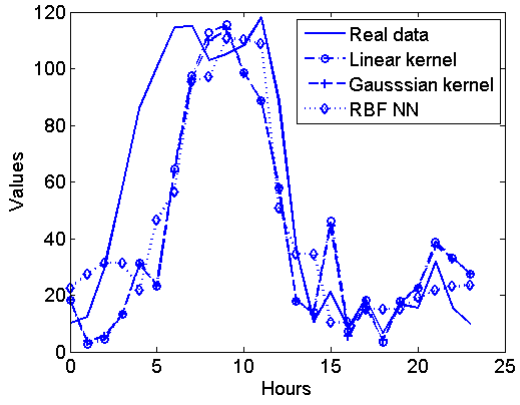


Figure 7: Prediction of levels of O_3 for 24 hours, for December 2005, for z=3

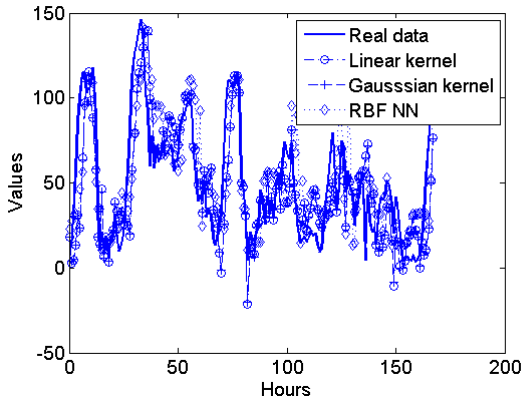


Figure 8: Prediction of levels of O_3 for one week, for December 2005, for z=3

The input data in December 2005 have big standard deviation. In this case due to the good generalization characteristics of the SVM models, the best prediction results for period of one week are achieved by the model built with SVM with polynomial kernel z = 1, 2, 3 and by SVM with Gaussian kernel for z = 4,..,8. The best results for prediction of 24 hours are achieved by model built with SVM with polynomial kernel for z = 1, 2, 5, 6, 7 and 8 and by SVM with Gaussian kernel for z = 3 and 4.

Figure 5 and 6 show the distribution of original O_3 data for the eleventh day and for one week in August, 2005. The same figures show the distribution of the predicted data that are obtained by the three models.

In August, when predicting the O_3 levels for one week, the best results are achieved by the model built with SVM with polynomial kernel for $z = 1, 3, 4, 6, 7$ and 8 and by SVM with Gaussian kernel for z = 2 and 5. The best results for prediction of 24 hours, the best results are achieved when using the model built with SVM with polynomial kernel for z = 1, 3, 4, 6 and 7 by SVM with Gaussian kernel for z = 2, 5 and 8.

6 Conclusion

The paper describes an attempt to predict the of hourly missing data for O_3 concentrations in the ambient air using SVM and RBF NN at the municipality Karposh III, in Skopje, Macedonia.

We developed a complete system for filling the gaps of missing hourly data by predicting the levels of O_3.

The built models for prediction of concentrations of ozone are examined for prediction of 8 consequent hourly values. The best results are achieved by the model built with SVM with polynomial kernel for prediction of 24 hours for December and August, 2005. In one case, the best results were achieved by the model built with SVM with Gaussian kernel, for prediction of one week for December, 2005. We should conclude that models built with SVM achieve better results than models built with RBF NN.

Finally we may conclude that SVM models give better results when predicting time series and they offer several advantages before the conventional RBF NN. In this paper we examined the free parameters of Gaussian kernel C, ϵ and σ and we conclude that only parameter σ has significant influence on the results from the offered models. Unlike the SVM models, the conventional RBF models parameters like the size of the network, the learning parameter and the training of the network play big role in the performances of the built model. Further on, are a result of the Structural Risk Minimization Principle, models built with SVM provide better prediction results compared with SVM models. Finally, using SVM we overcome the problems of neural networks like overfitting and local minima.

Although it is not possible to use the exact same models to predict the concentrations on the other measurement places in the country, still the presented methodology is general and it may be used for building new models for the other measurement places. The new models will be trained with data measured at the local measurement sites.

Models for prediction of ozone concentrations may be further extended. The developed model for ozone prediction uses data for NO_2, O_3, temperature and humidity. It may be extended with additional data for NOx, data for emissions from vehicles and other known sources of ozone. Similarly, the models may be extended with additional meteorological parameters.

The developed models are based on real data. In future, the presented methodology could be used for development of models that will take into consideration emissions from large combustion plants or the complexity of terrain where the prediction is performed. The missing data may be fulfilled with the built models, and after that the "new" data sets may be used for further prediction of concentrations of the same or other parameter. In the further research, it is possible to add the additional chemical or time dependence among the parameters , that will lead to new models for prediction.That way, in future, we may improve the use heuristic formula for prediction of ozone concentrations and decrees the MEA.

The experiments showed that the SVM is an appropriate tool for prediction of O_3 levels both for summer and winter seasons. The method gives good results and may be used by MoEPP for filling the data gaps for hourly O_3 values for short periods of time.

References

[1] A. Pelliccioni, T. Tirabassi. (2006) Air dispersion model and neural network: A new perspective for integrated models in the simulation of complex situations. s.l. : Environmental Modelling & Software, Vol. 21, pp. 539–546.

[2] Agirre-Basurko, E., Ibarra-Berastegi, G. and Madariaga, I. (2006) Regression and multilayer perceptron-based models to forecast hourly O3 and NO2 levels in the Bilbao area. *4,* Environmental Modelling and Software, Vol. 21.

[3] Kolehmainen M., Martikainen H., Ruuskanen J. () Neural networks and periodic components used in air quality forecasting. *815–825,* 2001 : Atmospheric Environment, Vol. 35.

[4] Wang W., Lu W., Wang X., Leung A.Y.T. (2003) Prediction of maximum daily ozone level using combined neural network and statistical characteristics. s.l. : Environment International, Vol. 1049, pp. 1–8.

[5] Smola, Alexander J. and Scöpholf, Bernhard. (1998) *A tutorial on Support Vector Regression, Statistics and Computing.*

[6] Vapnik, V., Golowich, S. and Smola, A. (1997) Support vector method for function estimation, regression estimation and signal processing. Cambridge : MIT Press, Neural information processing sytems.

[7] Canu, Stephane and Rakotomamonjy, Alian. (2001) Ozone peak and pollution forecasting using support vectors. *IFAC Workshop on environmental modeling.*

[8] Weizhen Lu, Wenjian Wang, Leung, A.Y.T.Siu-Ming Lo Yuen, R.K.K. Zongben Xu, and Huiyuan Fan. (2002) Air pollutant parameter forecasting using support vector machines. *IJCNN, ,* s.l. : I Neural Networks. 630–635.

[9] Wei-Zhen, Lu and Wen-Jian, Wang. (2005) Potential assessment of the "support vector machine" method in forecasting ambient air pollutant trends. s.l. : Elsevier, Chemosphere , Vol. 59, pp. 693-701.

[10] Cao, Lijuan. (2003) Support vector machines experts for time series Forecasting. s.l. : Elsevier, Neurocomputing, Vol. 31, pp. 321-339.

A Method for Calculating Acknowledged Project Effort Using a Quality Index

Marjan Heričko and Aleš Živkovič
University of Maribor, Faculty of Electrical Engineering and Computer Science
Smetanova 17, SI-2000 Maribor, Slovenia
E-mail: {marjan.hericko, ales.zivkovic}@uni-mb.si

Zoltán Porkoláb
Eötvös Loránd University, Faculty of Informatics
Pázmány Péter sétány 1/C , H - 1117 Budapest, Hungary
E-mail: gsd@elte.hu

Keywords: project effort, product metrics, software size, function points, object-oriented metrics

Received: August 11, 2007

Software size is the fundamental metric for project planning. Effort and duration are calculated based on the size estimate. However, for a given software size, the actual development effort could be significantly different. The question is whether the increase in effort is due to the low productivity of the development team or higher product quality. While higher product quality is highly desirable and usually worth investing in, the reasons for additional effort might be elsewhere. In the research presented in this paper, the focus is on the correlation between code quality and productivity. Code quality is only one aspect of product quality. This paper presents a method for calculating a new type of project effort named "acknowledged effort". Acknowledged effort is calculated based on the actual effort and code quality. This new type of effort reflects not only the project's size and the productivity of the development team, but also the quality aspect of the delivered software system.

Povzetek: V prispevku je analizrana korelacija med kakovostjo programske kode in produktivnostjo.

1 Introduction

Software size is an elementary measure often used to calculate project effort, costs, productivity and duration. In practice, the actual effort measured during the project development time could be significantly different although the estimated project size is the same. The effort is influenced by several factors like the complexity of the solution, development team size, development platform, etc. In this research, the focus is on the code quality that could as well influence the project total effort. The quality is defined as the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs. [10]. In terms of measures, it is a collection of metrics that cover categories like functional correctness, maintainability, efficiency, portability, usability and dependability [5,7,8,12]. The number of metrics used to determine product quality is not well defined and it could range from just a few to a hundred or more [2,12]. In contrast, there is the idea of a single number that express quality - the quality index (QI). The quality index is based on the 20/80 rule. According to some findings, 20% of variables can capture 80% of the intrinsic quality. The second principle behind the quality index is consistency and repeatability. If we perform the same procedure over and over again it will provide us with insight into product quality, regardless of its absolute accuracy in general applications. However, the standard deviation of the accuracy should be low in order to get valuable results. In this research, the code quality is measured in order to justify the deviations in project effort. Besides productivity, also code quality should be evaluated when comparing the performance of the development teams. The general functional relation between the productivity and software size is [1]:

$$P = \frac{E}{S} \qquad \text{(Eq. 1)}$$

where E is the actual effort spent developing some functionality and S is the total size of the functionality in question. For software size estimation, different methods are used [1,9,14,15], all of which have their roots in the Function Point Analysis (FPA) method.

The main contribution of the research presented in this paper could be summarized as:

1. The identification of the minimal object-oriented source code metrics set that might compose the quality index (QI) used together with software size in function points.
2. The definition of acknowledged effort that combines a team's productivity with the code quality. The actual project effort is then compared to the acknowledged effort in order to evaluate project results.

The research is based on the following assumptions and restrictions:

- the delivered code is complete and fully functional
- financial results, restrictions and influences are not considered in the proposed method
- the mean value for the productivity of some data set is valid for average code quality e.g. the projects are of different quality, however if the number of projects in the data set is big enough the data set would represent the average code quality.

This paper is divided into six sections. In the next section, the FPA method is briefly presented. The product metrics are introduced in section three. The main idea and the proposed method can be found in section four. The last section summarizes the findings related to a set of object-oriented projects and discusses the potential direction for future work.

2 Size Estimation for Object-Oriented Projects

Albrecht [1, 9] introduced the Function Point Analysis (FPA) method in 1979. Since then, it has become the most important method for software size estimation. The method introduced a specific way of representing a software system and distinguished between data functions and transactional functions. Data functions (DF) are further divided into internal and external logical files (ILF and EIF) assigning different weights to each data function type. The transactional functions (TF) describe functionality through three abstract types, namely: external inputs (EI), external outputs (EO) and external inquiries (EQ). To be able to determine the contribution of the FPA element (ILF, EIF, EI, EO or EQ) to the final estimated size value, the complexity is assigned to each element. The complexity is determined by the number of simple data elements named Data Element Type (DET) or structured elements named Record Element Types (RET). To get actual values in Function Points (FP), the tables defined in the method are used. The FPA abstraction concept is easily applied to structured analysis and design artefacts. The mapping of entities attributes and processes to FPA elements is straightforward.

The method was intended for all domains, although in practice, its accuracy is different within different domains. From a practical standpoint, it can be concluded that the FPA method application is more difficult with object-oriented projects. The elements and constructs of the FPA method are not directly applicable to object-oriented concepts also used within the Java and .NET development platforms. Therefore, a mapping of object-oriented concepts into FPA elements is needed. The mapping is not defined within the FPA method itself and is consequently not uniform. Different authors have proposed different mapping functions [3, 14, 15, 17], mostly in the form of additional rules. Information is gathered from different diagrams (e.g. Use Case diagrams, class diagrams, sequence diagrams)[14, 17] which are then considered separately. In one of our previous research, the OO-to-FPA mapping was defined and automated [17]. More detailed research of the FPA transformation tables has shown that the weight factors of the standard FPA method have to be calibrated for use in object-oriented projects [3, 4, 15].

3 Product Metrics for Object-Oriented Systems

In the software engineering community the term metric has been used in many distinct ways. For the purpose of this research, metrics are defined as a function, whose value is derived from a product, process or resource. It is important to distinguish between objective and subjective metrics. An objective metric is a function whose input is software data (elements) and whose output is a single numerical value. Subjective metrics, on the other hand, attempt to track less quantifiable data and usually depend on subjective judgment. When speaking about quality metrics the obtained metric value indicates the degree to which software possesses a given quality attribute. Therefore quality metrics are an indirect measure of software quality. We need validated metrics, metrics whose values have been proven to be statistically associated with corresponding software attributes. For object-oriented software the following metrics are often used [2, 13]:

- Weighted Methods per Class (WMC) - the sum of the complexities of the methods of a class (if all method's static complexities are considered to be unity, the number of methods). The number of methods and the complexity of methods involved are indicators of how much time and effort is required to develop and maintain the class. A large number of methods might limit the possibility of reuse since the class becomes too application specific.
- Depth of Inheritance Tree (DIT) - depth of the inheritance of the class. Inheritance through classes increases its efficiency by reducing the redundancy. However, the deeper inheritance hierarchy makes the behavior more difficult to predict and understand. There is no general threshold value for this metric. The threshold must be determined within the development team.
- Number Of Children (NOC) - the number of immediate sub-classes subordinated to a class in the class hierarchy. The greater the number of children in the inheritance hierarchy the greater the reuse. Then again a large number of children of a class might indicate improper abstraction for a parent class. In general the high DIT value and low NOC means better reusability but worse maintainability. It also has a negative impact on understandability and is more difficult to modify. Since there are no empirical or theoretical boundary values, the

developers should find the proper threshold value for the system under development.

- Response For a Class (RFC) - the sum of the number of its methods and the total of all other methods that they directly invoke. If the number of methods invoked in response to a message received by an object is large, the maintenance and testing are more demanding. Again there is no specific threshold value for the metrics.
- Coupling Between Objects (CBO) - the number of non-inheritance related couples with other classes (class is coupled with another if its methods use the attributes of the other class). The reusability of classes and/or subsystems is low when coupling between them is high, the system is also harder to understand. Normally a class should have a low coupling with the rest of the classes. A high coupling between different parts of a system has a negative impact on the modularity of the system and is usually a sign of poor design.
- Lack of Cohesion in Methods (LCOM) - the number of disjoint sets produced from the intersection of the set of attributes that are used by the methods reduced by the number of method pairs acting on at least one shared attribute. The LCOM metric is a value of the dissimilarity of the methods in the class. A high LCOM value in a class indicates that it might be a good idea to split the class into two or more sub classes. The metrics help identify flaws in the design of a program structure. The high LCOM values are associated with lower productivity, greater design and rework effort. The LCOM could be used as a predictor for maintenance effort.
- Method Hiding Factor (MHF) - sum of the invisibilities of all methods defined in all classes / total number of methods.
- Attribute Hiding Factor (AHF) - sum of the invisibilities of all attributes defined in all classes / total number of attributes
- Method Inheritance Factor (MIF) - sum of inherited methods / total number of available methods
- Attribute Inheritance Factor (AIF) - sum of inherited attributes / total number of available attributes
- Polymorphism Factor (POF) - actual number of possible different polymorphic situation / maximum number of possible distinct polymorphic situation
- Coupling Factor (COF) - actual number of couplings not imputable to inheritance / maximum possible number of couplings.
- Cyclomatic Complexity (CC) - in object-oriented design, the metrics represents the complexity of a method and indirectly also complexity of a class. The value should be as low as possible. The values between 10 and 20 are considered as an upper limit for metrics.
- Maintainability Index (MI) - predict the maintainability of the software combining several elementary metrics. Two versions are in use. The first version uses three elementary metrics to calculate the index and the second uses four metrics. The fourth metrics evaluates the average number of comments per class. However, it is not clear if the greater number of comments actually increases the ease of code maintenance. In this research the first version of the metrics will be used. The threshold values for MI are: MI<65 indicate poor maintainability, 65≤ MI ≤ 85 fair maintainability and MI > 85 promises excellent maintainability [2].

In addition to the described metrics, some size-related metrics should also be considered. Table 1 summarizes the candidate metrics classified according to the class/method level.

Table 1: Size Related Metrics

Class Level	Method Level
number of methods	number of parameters
number of properties	number of local variables
number of constructors	number of exception blocks
number of nested classes	max stack size
number of data fields	number of instructions
number of events	number of all operators in the method
number of attributes	number of distinct operators
the number of all instructions	number of operands

In order to collect and analyze metric data of object-oriented projects, the tool in the Microsoft .NET framework 2.0 was developed. The input is arbitrary executable format for the Microsoft platform. The parser that is a part of the tool performs an analysis directly on the common intermediate language code as defined in the .NET framework. In addition to the metrics presented in section three (not all metrics are supported in this version), the tool collects the data presented in Table 1. The tool is also described in [18].

Based on the collected data and the code quality metrics presented in section three, the correlation between different metrics were investigated as well as their potential impact on project effort and code quality. Based on the findings, the subset of code metrics was selected. The selected metrics that are used for calculating quality index (QI) and acknowledged effort (E_{ACK}) are listed in the next section.

4 The Proposed Method

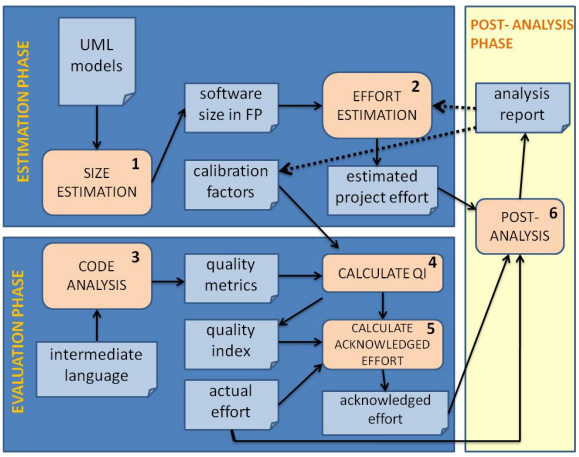


Figure 1: The Schematic View on the Proposed Method

Figure 1 shows the proposed method for calculating the acknowledged effort. In the analysis phase the project size and effort are calculated based on the UML models and projects characteristics. After the implementation phase the code analysis is performed. Based on the code analysis and actual effort the acknowledged effort is calculated. The acknowledged effort is defined as:

$$E_{ACK} = E_A * (RE_1 + (QI * RE_2)) \qquad \text{(Eq. 2)}$$

where E_{ACK} is the acknowledged effort expressed in hours, E_A is actual effort in hours, RE_1 and RE_2 are reward factors and QI is the quality index that has no unit. In our research the value for RE_1 is 0,7 and RE_2 is 0,1 which influences the actual effort for ±20 %.

The QI is defined as:

$$QI = \frac{\sum_{i=1}^{n} PMQR_i}{n} \qquad \text{(Eq. 3)}$$

$$PMQR \in \{1..5\}$$

$$PMQR_i = f_i(mv_i)$$

where $PMQR$ is the product metric quality rating, n is the number of code metrics used in the calculation, mv is a code metric value and f is the function that transforms metric value for metric i to the product metric quality rating. Code metrics that were considered in this research are described in section two. The quality rating transformation function is defined for each metric individualy. For the purpose of this research f is a step function that is defined based on the individual metric threshold values. An example for the maintability index (MI) metric is shown in Figure 2.

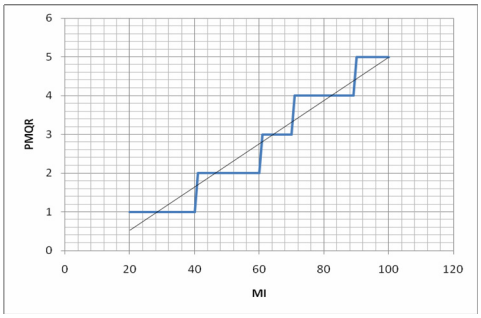


Figure 2: An example of the function f for the MI metric

QI is composed of n product metrics. The number of metrics and its type should be defined according to the project and environment characteristics. Each product metric has its threshold values. The threshold values are project specific and should be calibrated considering the folowing attributes:

- development team (experience level, team size, number of roles involved, etc.),
- development environment (platform, technology, process model, tools, etc.),
- domain (telecommunications, insurance, banking, etc.),
- customer (long term agreements, inhouse development),
- development type (off the shelf, research projects, critical systems, new development, reengineering, etc.).

In this research, the metrics described in section three were considered for selection. The narrowed list includes the following metrics:

- Depth of Inheritance Tree (DIT),
- Number of Children (NOC),
- Weighted Methods per Class (WMC),
- Coupling Between Objects (CBO),
- Response fo Classes (RFC),
- Lack of Cohesion in Methods (LCOM),
- Cyclomatic Complexity (CC) and
- Maintainability Index (MI)

The DIT and NOC make a complementary pair and should be considered as a pair [6, 7]. In this research only the DIT was used. WMC, CBO and RFC are highly correlated [2, 6]. The CBO was selected for the final set. The extended cyclomatic complexity (ECC) is included in the calculation of the MI which makes CC highly correlated to the MI values [6, 7]. Therefore the CC metric is also excluded from our metrics set. Thus the final metrics set used for calculating QI consists of DIT, CBO, LCOM and MI.

For the product metrics (PM) in the final metrics set, four functions f that transform metric values to the product metric quality ratings (PMQR) were defined. The PMQR range is one to five and the range of the PM is metric specific. Figure 3 presents the transformation function for all four metrics (DIT, CBO, LCOM and MI). Please note that these step functions should be calibrated before their use in a different environment.

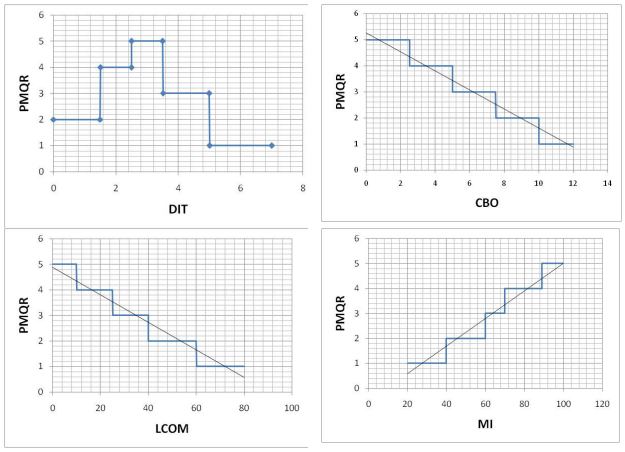


Figure 3: Step function f for the DIT, CBO, LCOM and MI metrics

Calculating acknowledged effort is only one possibility for applying the quality index (QI) defined in this paper. Another possibility is to calculate the corrected productivity of the development team. The corrected productivity is defined as the normalized productivity for the delivered results. In case of bad design and low quality code, the productivity calculated from the effort and size -- sometimes called actual productivity -- is higher then the corrected productivity defined here. If the delivered code is of outstanding quality the actual productivity is lower than the corrected productivity calculated using the quality index.

$$P_A = \frac{E_A}{S} \qquad \text{(Eq. 4)}$$

$$P_C = P_A * (RE_1 + (QI * RE_2)) = \frac{E_{ACK}}{S}$$

where P_A is actual productivity for the current project calculated from the actual effort E_A (sometimes also called recorded effort). P_A is calculated at the end of the project. P_C is corrected productivity, RE_1 and RE_2 are reward factors, S is software size and QI is quality index.

The proposed method was used on a set of OO projects in order to explore the acknowledged effort on real projects. Table 2 summarizes the metrics values for three groups of projects. In the first group are smaller student projects written in Java. The students were from the last grade of the computer science study program. Since Java is already introduced in the first year and used throughout the study program for individual and/or group projects at different subjects it can safely be assumed that in the last year, students have good programming skills and sufficient development experience. In the second group are industry projects developed on the Microsoft .NET platform. The development team was experienced and used sophisticated development approaches like custom code generators and design patterns. The third group is a control group. The projects are for well known products, developed by highly experienced development teams. The metrics data is from the master thesis prepared at the Uppsala University in Sweden [2]. The DIT, NOC, CBO and LCOM metrics were collected on the class level. Therefore two values are provided in the table. The first value is the mean and the second is the standard deviation. MI is calculated at the project level, thus only one value is in the table.

Table 2: The Values for Selected Code Quality Metrics

	DIT	*CBO*	*LCOM*	*MI*
GROUP 1				
Project 1.1	3,35 2,47	11,82 11,23	95,56 13,33	65,19
Project 1.2	1,07 0,38	4,37 5,50	65,19 34,25	49,64
Project 1.3	2,91 2,47	12,55 12,74	68,14 34,54	34,48
Project 1.4	1,82 1,85	5,26 8,49	71,18 28,95	30,31
Project 1.5	2,57 2,18	11,07 12,45	73,17 28,52	34,29
Project 1.6	1,54 1,50	5,7 9,62	72,71 35,41	48,36
Project 1.7	1,00 0	4,20 3,52	51,25 43,18	29,86
Project 1.8	1,21 1,26	2,79 5,36	67,20 35,25	45,37
Project 1.9	4,7 2,13	17,48 12,16	73,91 26,85	72,34
Project 1.10	2,85 2,03	9,00 12,23	66,20 25,08	68,57
Project 1.11	0,36 0,48	1,31 2,57	85,00 10,68	87,49
Project 1.12	0,75 1,68	2,25 5,10	15,63 30,15	49,53
Project 1.13	2,00 2,03	8,58 8,28	74,78 30,47	80,34
Project 1.14	3,64 2,51	14,49 12,42	85,65 7,10	77,04
GROUP 2				
Project 2.1	2,47 1,01	0,79 4,08	4,43 18,04	87,53
Project 2.2	2,23 1,11	5,96 8,81	31,06 37,42	79,98
Project 2.3	3,28 1,71	9,00 8,90	23,57 40,25	89,59
Project 2.4	1,20 0,44	5,00 11,18	12,40 27,72	77,41
GROUP 3				
Project 3.1	0,58 0,75	8,36 5,87	0,33 0,33	129
Project 3.2	0,21 0,41	9,47 6,18	0,41 0,36	164
Project 3.3	1,23 0,83	6,53 4,56	0,38 0,19	182

Table 3 presents data for size, effort and *QI*. Project size is expressed in function points as well as in lines of code (LOC). In column four is the actual effort E_A in hours followed by the acknowledged effort E_{ACK}. In the last column are the values for the quality index calculated following the proposed method. For most of the student projects the *QI* is less than three. Consequently the acknowledged effort is smaller than the actual effort reported by the students. The projects in the second group demonstrate better code quality then is normally expected (*QI*>3,0), the acknowledged effort is higher.

The projects from the third group are not included in the table since the actual effort for them is unknown.

Table 3: Size, Effort and QI Results for Test Projects

	Size (FP)	Size (LOC)	E_A (h)	E_{ACK} (h)	QI
GROUP 1					
Project 1.1	72	4.216	95	90	2,50
Project 1.2	65	4.176	105	97	2,25
Project 1.3	163	1.760	191	172	2,00
Project 1.4	37	2.006	88	81	2,25
Project 1.5	88	2.777	192	173	2,00
Project 1.6	157	3.642	57	54	2,50
Project 1.7	71	1.782	171	158	2,25
Project 1.8	173	2.159	54	50	2,25
Project 1.9	110	3.400	143	132	2,25
Project 1.10	35	1.686	45	44	2,75
Project 1.11	43	2.576	210	210	3,00
Project 1.12	35	623	39	40	3,25
Project 1.13	60	985	125	122	2,75
Project 1.14	70	4.189	165	153	2,25
GROUP 2					
Project 2.1	2.122	93.978	8.800	10.120	4,5
Project 2.2	440	32.532	1.067	1.120	3,5
Project 2.3	1.987	156.122	2.133	2.347	4,0
Project 2.4	13	771	56	59	3,5

5 Conclusion

The typical evaluation of completed software projects includes costs, effort and completeness of the delivered functionality. In this research the focus was only on project effort. From the management point of view, the recorded effort is not necessarily the acceptable project effort when taking into consideration the quality of the delivered code. In this paper, the idea of acknowledged effort was presented. Acknowledged effort combines actual effort with a quality index. The quality index is a single value that represents the quality of the delivered code. The management could then reward or penalize the development team for arbitrary percentages in accordance with the code quality. The formula provided in this paper should be calibrated accordingly. The idea presented in the paper was tested on the sample data set, including 18 projects. The results demonstrate when the effort should be smaller than the actual effort as well as when the quality of code is better than average and the developers should be additionally rewarded for their work.

In the future, the method will be tested with different metrics sets and additional project in order to validate the sensitivity of the proposed method.

Acknowledgement

This work was realised as a part of the Slovenian-Hungarian bilateral research project BI-HU/06-07-01 Application of OO Metrics to estimate .NET project software size.

References

[1] Albrecht,A., (1979). Measuring Application Development Productivity, *IBM Applications Development Symposium*, pp. 83-92.

[2] Andersson, M., Vestergren, P., (2004). *Object-Oriented Design Quality Metrics* (Master Theses), Uppsala University.

[3] Antoniol,G., Lokan,C., Caldiera,G., and Fiutem,R., (1999). A Function Point-Like Measure for Object-Oriented Software, *Empirical Software Engineering*, Springer, pp. 263-287.

[4] Antoniol,G., Fiutem,R., and Lokan,C., (2003). Object-oriented function points: An empirical validation, *Empirical Software Engineering*, Springer, pp. 225-254.

[5] Arisholm, E., (2006). Empirical assesment of the impact of structural properties on the changeability of object-oriented software, Information and Software Technology, Elsevier. pp. 1046-1055.

[6] Chidamber, S.R., Darcy, D. P., Kemerer, C.F. (1998). Managerial Use of Metrics for Object-Oriented Software: An Exploratory Analysis, *IEEE Transaction on Software Engineering*, IEEE, pp. 629-639.

[7] Fioravanti, F., Nesi, P., Stortoni, F., (1999). Metrics for Controlling Effort During Adaptive Maintenance of Object Oriented Systems, *Proceedings of the IEEE International Conference on Software Maintenance*, IEEE, pp. 483-493.

[8] Gyimóthy, T, Ferenc, R., Siket, I., (2005). Empirical Validation of Object-Oriented Metrics on Open Source Software for Fault Prediction, *IEEE Transaction on Software Engineering*, IEEE, pp.897-910

[9] IFPUG, (2004). *Function Point Counting Practices Manual, Release 4.2*, International Function Point Users Group.

[10] ISO, (1986). *ISO 8402 - Quality management and quality assurance*, ISO.

[11] ISBSG, (2001). *Practical Project Estimation, A toolkit for estimating software development effort and duration*. International Software Benchmarking Standards Group.

[12] Marinescu, R., Ratiu, D., (2004). Quantifying the Quality of Object-Oriented Design: the Factor-Strategy Model, *Proceedings on the 11th Working Conference on Reverse Engineering (WCRE'04)*, IEEE.

[13] NASA, (1995). Software Quality Metrics for Object Oriented System Environments, Mational Aeronautics and Space Administration.

[14] Uemura,T., Kusumoto,S., and Inoue,K., 2001. Function-point analysis using design specifications based on the Unified Modelling Language. *Journal of Software Maintenance and Evolution-Research and Practice*, Interscience, pp. 223-243.

[15] Živkovič,A., Heričko,M., and Kralj,T., (2003). Empirical assessment of methods for software size estimation. *Informatica (Ljubljana)*, Slovenian Society Informatika, pp. 425-432.

[16] Živkovič,A., Heričko,M., Brumen B., Beloglavec S., Rozman I., (2005a). The Impact of Details in the Class Diagram on Software Size Estimation, *Informatica (Lithuania)*, Institute of Mathematics and Informatics, pp. 295-312.

[17] Živkovič, A., Rozman, I., Heričko, M., (2005b). Automated Software Size Estimation based on Function Points using UML Models, *Information & Software Technology*, Elsevier, 881 - 890

[18] Živkovič, A., Heričko, M., Porkoláb, Z., (2007). Evaluating the Correlation Between Code Quality, Software Size and Effort, *Proceedings of the 10th International Multiconference Information Society IS 2007 Volume A*, IJS.

An Approach to Knowledge Transfer in Software Measurement

Jari Soini
Tampere University of Technology, Pohjoisranta 11, P.O.Box 300, FI-28101 Pori, Finland
E-mail: jari.o.soini@tut.fi, www.pori.tut.fi

Keywords: software engineering, software measurement, knowledge transfer

Received: June 14, 2007

Being successful in knowledge management is one of the most essential factors in any organization if one wishes to succeed in business in the future. This aspect is emphasized particularly in knowledge intensive organizations, such as software companies. Existing knowledge must be easily available for all and it must be possible to bring new information to the attention of everyone. This article introduces one example of a solution implemented to deal with this challenge. The case study described in this article deals with software engineering measurement and related knowledge collection, distribution and also utilization in practice. This example is one approach to trying to solve the challenges related to both the growing needs for effective knowledge transfer and for enhancing the utilization of measurement in software engineering.

Povzetek: Prispevek obravnava meritve programske opreme glede prenosa znanja.

1 Introduction

In a networking economy, which is typical of today's information society, the ability of organizations and their members to cooperate, interact and share their information and knowledge is a prerequisite for strategic operation [1],[2],[3],[4]. Many studies support the view that collaboration and knowledge sharing between companies are effective means to enhance profit and growth [5],[6],[7],[8]. Nowadays, the success of an organization depends, maybe more than ever, on its ability to create and share knowledge effectively and efficiently [9]. There is a strong belief that the systematic transformation of human capital into value requires structural capital as a multiplier, to realize a sustainable earnings potential for the organization [10],[11],[12],[13]. This has forced companies to build a co-operative relationship with other organizations and try to increase their own learning and knowledge by utilizing the knowledge sharing which occurs during this interaction. However, successful networking and collaboration with other companies demands great openness, which is not easy for every company [14],[15]. This article gives us an empirical example of how to implement and manage this in practice.

The research topic of the case study relates to software engineering measurement. Measurement is one of the key elements in receiving feedback and evaluating the processes used in the organization. In software engineering, as knowledge of measurement theory and experience increase, measurement has begun to be perceived as an effective method of understanding, controlling, steering, predicting and improving software development and maintenance projects [16]. Fenton and Pfleeger [17] defined software measurement as the continuous process of defining, collecting, and analysing data on the software development process and its products to understand, control and optimize the process and its products. This definition describes quite exhaustively what software measurement is all about. Nowadays software measurement is one of the key components in an organization's ability to maintain their competitiveness in a rapidly changing business environment [18]. There is a growing need for the use of objective information in decision making, and appropriate measurement enables all levels in software organizations to obtain this kind of objective information [18]. Organizations need a deeper understanding of their own processes and to do this, they need measurement data on their current processes. Although the advantages of measurement in the software development process are indisputable, the popularity of use of measurement methods in practice is rather limited [18],[19],[20]. Implementing measurement in software engineering raises many challenges. Very often difficulties arise when trying to focus the measurement. In many cases it is unclear what should be measured and also how the measurement data obtained should be interpreted [21],[22]**.** Choosing the correct measurement entities and ranking the importance of measurement indicators is a challenging task [23],[24],[25].

A research project was established in autumn 2005 with the aim of examining and trying to find potential procedures and techniques to solve problems related to product and process measurement in software engineering. This research was carried out in close collaboration between two universities and Finnish software organizations. One of the most essential sources of information in this work was the current measurement practices and metrics used in software organizations. Determining a method and developing an instrument for capturing, analyzing and sharing this measurement

knowledge between individuals and organizations was one of the central parts of the research. This article describes the instrument developed, *the information system*, for enhancing and transferring knowledge, related to software engineering measurement. The main contribution of this research was to give an example of how knowledge transfer in software engineering measurement can be organized and managed in practice.

2 Background of the study

2.1 Principles and goals of knowledge management

The starting point of the research was to understand the basis of both knowledge management and software engineering measurement. With regard to knowledge transfer we must first understand the basis of knowledge management. Figure 1 describes the theoretical framework of knowledge management [26]. As we can see, knowledge management (KM) refers to the activities involved in discovering, capturing, sharing, and applying knowledge. KM processes are meant to assist operations. Furthermore, processes are supported by KM systems, which are the integration of technologies and mechanisms. KM sub-processes (such as combination, socialization, externalization, internalization, exchange, direction, and routines) facilitate the broad processes and KM systems themselves rely on a current KM infrastructure.

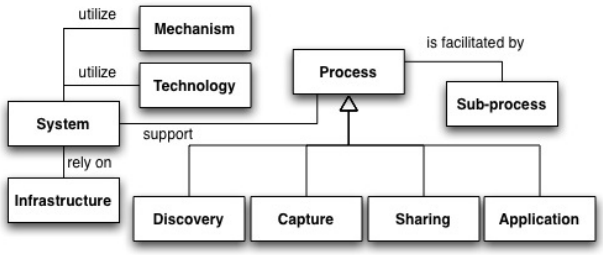


Figure 1: The concepts of knowledge management [26].

In general, knowledge management can be seen as a matter of improving conducive ways of thinking, practices and developing support systems to promote knowledge sharing. Success in knowledge transfer, which is one of the key issues in this study, is one prerequisite for creating new knowledge and organizational learning. The company must create a co-operative relationship with other organizations and try to increase their own learning and knowledge by utilizing the knowledge sharing and transfer which occurs during this intercourse [27]. The overall aim is to reduce the uncertainty of the operational environment by ensuring that the company has the opportunity to access wider knowledge of the business environment. Knowledge sharing, which is a consequence of this kind of collaboration, generates additional value for every organization, in this case by increasing its knowledge capital, thereby improving its competitiveness.

2.2 A case study – the SoMe project

The motivation for the research derived from issues observed in relation to the software process and product quality, at national level [28] as well as international level [29],[30],[31]**.** The widely accepted assumption is that software quality, in general, is caused by and dependent on the quality of the software development process. The accepted opinion is that most problems in software quality are based precisely on problems in the software development process [32],[33],[34]. In practice, it has proved difficult to define the key functional process and product measurements and many software companies have found measurement to be challenging and problematic [35],[36]. To promote a better understanding of measurement and to offer a robust and pre-selected set of metrics suitable for different kinds of business goals, FiSMA (the Finnish Software Measurement Association) [37], initiated the SoMe (Software Measurement) project in autumn 2005 together with Tampere University of Technology (TUT) [38] and the University of Joensuu (UJ) [39]. FiSMA itself is a non-profit making organization created to promote the usage and utilization of software measurement to improve the quality of processes and products. Its members, who are also the participants of this study, consist of nearly 40 Finnish software companies, plus several universities and other public organizations. In the context of the SoMe project, different instruments and practices were studied to help solve the measurement problems related to the quality of both software process development and software products. The SoMe project focused especially on experience-based measurement data. Therefore, the sample was a set of software companies (a total of 10) who perform process measurement in practice. Nine of the participant companies can be classified as small and medium size companies (SMEs) [40], from the viewpoint of organizational units involved in software engineering (table 1). The common characteristics, shared by these companies, are that their core business is the supply of software projects and they carry out software development independently.

Company	A	B	C	D	E	F	G	H	I	J
Empl., total	195	200	200	220	280	450	1,200	3,200	15,000	24,000
Empl. In SE	195	30	200	200	150	35	30	120	5000	200

Table 1: Size of participating companies by the number of employees

The main idea and the ultimate objective of the SoMe project was to develop, in cooperation with the participants, a common and open information system for Finnish software companies to help monitor and measure the quality of their software processes and products. The final outcome of the SoMe project was an information system implemented in a web environment based on a large metrics database. The final database consists of three different types of measurement information (practical experience, literature and standards). The

developed information system utilizes a web-based repository of best practices.

3 Developing an instrument for transferring software measurement information

3.1 Method used for capturing, organizing and evaluating the collected knowledge

One of the main targets of the project was to get as much experience-based information in the database as possible. Therefore, the empirical part of the research was based on a series of interviews and questionnaires, created to collect the experiences of the companies about individual metrics and measurement practices in general. The research method used was to conduct interviews to address the research questions and the target group was quality managers. Information on the metrics used and current measurement practices were collected on a spreadsheet-style form, providing the basic data for this study (see Appendix A). The aim was to give as explicit a description as possible of all the process metrics used in the participating companies. The information on the current metrics (Appendix A) was collected before the interviews, because it provided the possibility to expand on it during the interview session. Figure 2 describes the process of how the knowledge was captured, modified, evaluated and shared.

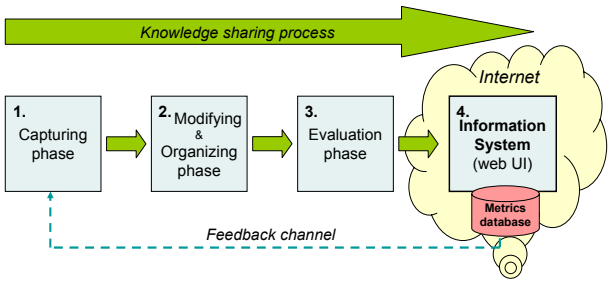


Figure 2: The phases for capturing and sharing knowledge [41].

After the capturing phase, the collected information was combined and organized by the researchers. The information captured was pre-evaluated and analyzed as to its suitability, usability and the correctness of the examined topic. The captured knowledge was also modified in the same framework. The metrics database consists of individual items of information, *knowledge items,* and the manifestations of these items are *metric documents.* The formula for the title level and the terminology used in all documents is congruent. This solution helps the end user to read, perceive the logic and make comparisons between knowledge items.

Before distributing the captured and modified knowledge via the information system, the applicability and intelligibility of the information together was evaluated with representatives of the participating companies (see ref. [41]). A support group was established inside FiSMA. The aim of this practice was to evaluate the knowledge collected with the end users before placing it in the information system and delivering it to the organizations. All the companies involved were able to take part in regular support group meetings. Other publications in relation to the SoMe project [41],[42] describe in more detail the process used for capturing, modifying, evaluating and distributing the measurement knowledge via the information system that was developed.

3.2 Overview of the information system

The final outcome of the SoMe project was a measurement knowledge base consisting of a large metrics database. The expression of this final outcome is a web-based information system for measurement knowledge transition, which was successfully implemented in April 2007. This information system is meant for individuals and organizations seeking appropriate software metrics and measurement practices for their needs. It offers a bi-directional link, from processes to metrics, and also vice versa. With the metrics offered, the information system enables organizations to utilize measurement for controlling and monitoring their software processes and products and thus enhances the quality of the software produced.

a. Ideology of the system

The aim of the SoMe project was to develop an information system which will enable organizations to control and improve their software development process and product quality. In relation to process quality, process assessment models like the ISO/IEC 15504 process assessment model (SPICE) [43] and Capability Maturity Model Integrated (CMMI) [44] are available, which allow evaluation of the quality of the current software development process in the organization. For process improvement work, organizations need a deeper understanding of their own processes and to do this they need measurement data [16]. With this measurement information they can reliably seek and find improvement objects in their processes [45],[46],[47]. This approach was selected as the starting point of the development work and steered the work throughout system implementation. This aspect also guides the search taxonomy design of the information system. The selected search taxonomy was created based on the CMMI and SPICE process assessment models (see Figure 3). The aim of this selection was to for the organizations to familiarize themselves with and utilize these assessment models in their operations. This aspect is crucial for the process improvement viewpoint. Assessing the actual state of the current processes, which is also the first step of process improvement actions, is one of the most typical purposes of the use of measurement in software engineering [23]. The main ideas for developing the information system were precisely to help organizations utilize measurement knowledge to control their software development process and also to use measurement for

supporting their process improvement work as well as being the source of objective information for management in their decision making.

b. Metrics database

The system itself works on a database, which contains information about software measurement literature, standards and actual metrics and measurement practices used in software organizations. The metrics database consists of individual items of information, *knowledge items* (individual metrics). A standard form, *a metric document*, is used for presenting each knowledge item. The formula for the title level and the terminology used in all metric documents is congruent with the others. This solution helps the end user to read, perceive the logic and make a comparison between the metrics. All the process and product metrics in the database, carefully analyzed relevant metrics collected from the participants (85 metrics) as well as additional relevant metrics found in literature or standards (22 metrics), have been modified using the same standard form.

The following knowledge exists on each individual metric captured in the system: Purpose, Formula (if required), Values (with a possible threshold value), Usage, Workloads (establishing the metric, collecting the data, using the metric), Risks, Experiences plus other information and References. All this information was collected with the spreadsheet-style form (Appendix A), If required, in experienced-base metrics, this information was supplemented and clarified during the personal interview sessions. In addition to this, in the standard form, the metric links to the search taxonomy can be seen (the processes that the particular metric relates to are shown). As an example, there follows a description of one metric; *Distribution of customer work*, and its information under the *Usage* title: "It is a derived metric used mainly by upper management to monitor how the maintenance work is distributed between corrective, adaptive and perfective work. All maintenance work must be classified (at least) into these categories and recorded in the time tracking system. To calculate a ratio, the formula is: work hours in one category divided by the sum of all maintenance work hours. To use the metric, a detailed time tracking system must be in use, where each employee records his/her working hours. Once a month the quality manager collects the totals of customer work hours and presents them as a table or a graph. Inserting data in the time tracking system requires manual work, but after that the results can be calculated automatically".

Appendix B describes the experience-based measurement objects collected by the participating companies. The ISO/IEC 15504 (SPICE) standard was used as a framework to classify the metrics into suitable categories named after the SPICE processes. An evaluation is also presented of the given characteristics (see Appendix A) in relation to the metrics used . A more detailed analysis of the results of these user evaluations is presented in a previous research paper [48]. The experience-based metrics are collected from companies whose capability levels varied between 2 and 3. The experience-based metrics inside the database can be utilized mostly at a SPICE capability level of 2 (56 %), at level 3 (35 %), and 9 % at level 4 (see ref. [41]).

c. Web-based information system

In the information system, the knowledge items (individual metrics) are linked to the process groups inside the assessment models. Every knowledge item also includes information for all the process groups to which it relates and in practice is linked. This characteristic enables the user to see the dependence between process groups from a metric viewpoint and gives important information when planning measurement activities (e.g. a measurement program). These connections are also seen from the process group viewpoint, as the proper metrics depend on the selected process group in the selected assessment model (SPICE or CMMI). This realization method enhances awareness of the relationship between process assessment and process measurement. As an example, Figure 3 presents the results of one search; SPICE assessment model / (Engineering process group) ENG 10 System testing. After making the selection, the results (individual metrics) of this search appear in the information system display (see right of the figure). The user can see directly all the related metrics and also a brief description such as: the name of the metric(s), a short summary of each metric and workload evaluation for establishing, collecting and using the metric. Depending on the given search selection, the system retrieves the particular metrics from the database that are linked to the selection (individual process group inside the assessment model). Selecting a particular metric (by clicking on the metric name field) calls up the detailed information of this metric (see example in section *b. Metric database*). Every user can examine and compare which metric or metrics are appropriate and also compatible with the needs and maturity of their organization. It is possible to make a new search with a new search criterion as many times as you want. Organizations decide and select an appropriate set of metrics according to their needs independently.

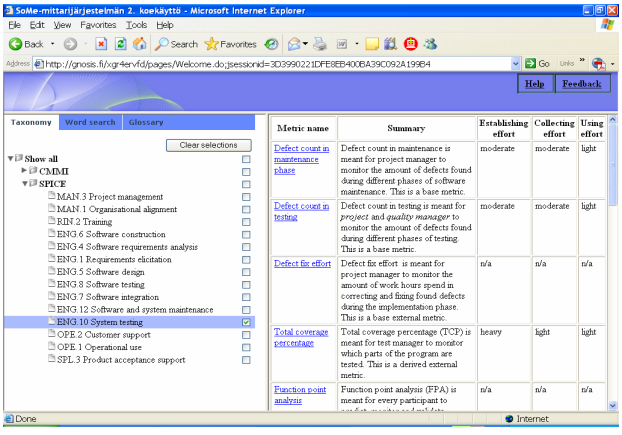


Figure 3: User interface (UI) of the web-based information system

The information system includes also a *word search* alternative (see Figure 3) for searching for a suitable metric. This feature was included because there may be organizations that are not familiar enough with the process assessment models to start using a system based on them. The starting point for planning the system web-user interface (UI) was that it must display as well as operate in such a simple way that the UI does not become an obstacle to the use of the system. A clearly and simply defined UI, both structurally and visually, is one of the most important factors when introducing new technical tools [49]. In this work, a lot of co-operation was made with the participant organizations and the people there were assumed to be the end users of the information system. This method is perceived as a functional approach when developing new applications for a certain target group [50].

Certain interactive activities are also involved in the information system. There are several reasons for this. Firstly, as stated above, the environment in software engineering is changing rapidly [51]. Therefore there could be a need for adding, modifying or even deleting some information related to the existing knowledge item (individual metric) in the system. Also, during the operation of the system, new experiences may arise and this interactivity allows new information to be added and also combined with the current item. This feature creates a line of communication between individuals and organizations, allowing them to share knowledge and learn from each other. Additionally, the information system includes a library. A glossary (see Figure 3) helps the user if the terms or concepts used in relation to measurement and metrics are not familiar. The terms and concepts used are mostly based on the terms and definitions used in software standards [52],[53],[54],[55]. This selection will guide the organization towards harmonized use of the terms related to software measurement.

4 Observations and evaluation of the developed system

4.1 General observations

The results of the research worth considering and evaluating are the research method used and the information system that was developed. The analysis and discussion presented here highlight some general observations on the research. The method used proved very useful for collecting and also evaluating information. The attitude and motivation of the participants to share their knowledge seemed very positive and the empirical data received met the requirements of the research. The method of data collection (a data form combined with interviews) and evaluating (with the support group) the data adds to its quality (see ref. [41]). The starting point and aim was to obtain more qualitative than quantitative data. The selected method supported the set goal very well. Overall, the method used during the research, as well as the created web-based information system for executing knowledge management and sharing, seem to be workable and seem to have achieved the goals which were set. The decision to connect the participants in the study from the very outset and the close co-operation with them throughout the project proved to be a good choice. This enhanced the knowledge-sharing process itself and also ensured the appropriateness of the final outcome. This method provides important information and feedback also for researchers on how to develop the system correctly and also maintained motivation for the researchers during a long project.

4.2 Evaluation of the information system

Below are evaluated briefly the main strengths and also some deficiencies of the information system.

Firstly, the strengths: It contains detailed information, mostly experienced-based, on the metrics involved in the system. With this information, organizations can initiate, or confirm and improve their existing measurement system. This kind of information seems to interest organizations and is very relevant especially if the organization wants or is intending to establish a measurement system. The most significant contribution from the technical point of view is a bidirectional link between processes and metrics. This feature helps and advises the users to identify the relevant metrics for controlling the particular process. The system guide advises the user on how to use measurement as an instrument for software process improvement. The user has the opportunity to see the connection and the dependence between software metrics and the software process groups. Becoming aware of these correlations is very important when planning, implementing or controlling the software processes or their improvement. In comparison with the other existing related tools, such as the knowledge PLAN and FP workbench by Software Productivity Research (SPR) [57] or the International Software Benchmarking Standards Group (ISBSG) database [58], the advantages of the new system are a closely specified description of the metrics themselves, a guiding feature for searching for the proper metric(s) and also an emphasis on software engineering measurement. With regard to the SPR (including information on over 2000 projects) or ISBSG (over 8000 projects), one disadvantage is the amount of purely numerical data. These systems are mainly intended for effort estimation and scheduling and for performing benchmarking. The SPR and ISBSG systems are more focused on a project viewpoint, rather than the process and product. One obvious strength of the new information system is also the congruent terminology of all the documents. This helps and advises the end user on how to read, perceive the logic and make a comparison between the knowledge items.

Secondly, however there are some deficiencies: the system is only to serve the issues related to the software process and product measurement; other subjects are excluded. This factor limits the utilization of the system. Also, the heavy emphasis on the software process

assessment model which has been selected may be a factor of uncertainty. This approach could be unfamiliar to some of the users. Related to the users' evaluation of the requested characteristics of the metric used, it must be noted that the answers are subjective. This may cause a bias in the evaluation results if they are regarded purely from a scientific point of view [59],[60]. It is also noteworthy that the experience data has been collected in organizations whose capability levels vary between 2 and 3. This fact must be taken into account when planning to utilize these experienced-based metrics.

Finally, some suggestions for future work. It may be useful to add some alternatives for the current search taxonomy (SPICE and CMMI) into the information system. For example, factors related to the organization, such as size and capability level, could be pre-selection factors when starting to seek the proper metrics. More feedback is needed when decisions are made relating to the improvement actions of the system. From a software measurement viewpoint, one relevant topic for future research could be to examine how the capability or maturity level of the organization may affect the utilization of the information system and how, and if, the required measurement information varies depending on the different levels. From a knowledge transition viewpoint, and also to validate the method developed and the tool itself, it could be interesting to implement this method for some other research subject in software engineering.

5 Summary

This article deals with the research based on issues observed in relation to software process quality. In order to examine the solutions, the research project (SoMe) was established, with the aim of studying how measurement can be utilized for solving this issue. The goal was to collect the relevant and experience-based measurement data for this purpose, and also to create an appropriate information system for delivering the received knowledge to the software organizations. This article describes the method used and gives an overview of the ideology and the structure of the developed web-based information system as well as the measurement data included it. The generated system offers information about different metrics and their applicability to measuring different processes. It also gives an example of how issues related to information can be collected and one approach to solving its transition.

References

[1] P. Drucker (1999). Knowledge-Worker Productivity. The Biggest Challenge. *California Management Review*, Vol. 41, no. 2, pp. 79-94.

[2] J. C. Jarillo and H. Stevenson (1991). Co-operative strategies – The payoff and the pitfalls. *Long Range Planning,* Vol. 24, pp. 64-70.

[3] M. Iivonen and M. L. Huotari (2004). *Trust in knowledge management and systems in organizations*. Idea Group Publishing, London.

[4] P. Sydänmaanlakka (2001). *Älykäs organisaatio. Tiedon, osaamisen ja suorituksen johtaminen*. Kauppakaari, Helsinki.

[5] H. Hinterhuber and B. M. Levin (1994). Strategic networks – The organization of the future. *Long Range Planning,* Vol. 27, pp. 43-53.

[6] C. Snow, R. Miles and H. Coleman (1992). Managing 21st century network organizations. *Organizational Dynamics*, Vol. 20, no. 3, pp. 5-20.

[7] J. S. Whipple, J. J. Gentry (2000). A network comparison of alliance motives and achievements. *Journal of business and industrial marketing,* Vol. 15, no. 5, pp. 301-322.

[8] I. Kulmala, A. Vahteristo and E. Uusi-Rauva (2002). *Ohjelmistoyritysten verkostoituminen – tutkimus suomalaisten ohjelmistoyritysten toimintatavoista ja tavoitteista.* Research Report 3/2002. Tampere University of Technology, Tampere, Finland.

[9] T. H. Davenport and L. Prusak (1998). *Working knowledge: How organizations manage what they know*. Harvard Business School Press, Boston, MA.

[10] L. Edvinsson, B. Kitts and T. Beding (2000). The next generation of IC measurement: The digital IC landscape. *Journal of Intellectual Capital,* Vol. 1, no. 3, pp. 263-272.

[11] M. R. Testa (2002). A Model for organization-based 360 degree leadership assessment. *Leadership & Organization Development Journal,* Vol. 23, no. 5, pp. 260-268.

[12] U. S. Bititci, V. Martinez, P. Albores and J. Parung (2004). Creating and managing value in collaborative networks. *International Journal of Physical Distribution and Logistics Management,* Vol. 40, no. 3, pp. 251-268.

[13] U. S. Bititci, V. Martinez, P. Albores and J. Parung (2004). Creating and managing value in collaborative networks. *International Journal of Physical Distribution and Logistics Management,* Vol. 40, no. 3, pp. 251-268.

[14] W. Seal, J. Cullen, A. Dunlop, T. Berry and M. Ahmed (1999). Enacting a European Supply Chain: A Case Study on the Role of Management Accounting. *Management Accounting Research,* Vol. 10, no. 3, pp. 303-322.

[15] M. Beeby and C. Booth (2000). Networks and inter-organizational learning: a critical review. *The Learning Organization,* Vol. 2, pp. 75-88.

[16] L. Briand, C. Differding and D. Rombach (1997). Practical Guidelines for Measurement-Based Process Improvement. *Software Process – Improvement and Practice,* Vol. 2, no. 4, pp. 253-280.

[17] N. E. Fenton and S. L. Pfleeger (1997). *Software Metrics: A Rigorous & Practical Approach, 2. Edition.* PWS Publishing Company, Boston, MA.

[18] J. McGarry, D. Card, C. Jones, B. Layman, E. Clark, J. Dean and F. Hall (2002). *Practical*

Software Measurement. Objective Information for Decision Makers. Addison-Wesley, New York.

[19] T. Varkoi (1999). Development of Measurement Programs to Support Process Improvement in Small Software Companies. *Proceedings of the FESMA'99 Conference.* Amsterdam, Netherlands, pp. 141-149.

[20] Software Engineering Institute (SEI) (2006). *The State of Software Measurement Practice: Results of 2006 Survey.* Technical Report, CMU/SEI-2006-TR-009, ESC-TR-2006-009.

[21] N. Habra, A. Abran and M. Lopez (2004). *Towards a framework for Measurement Lifecycle*. University of Namur, Technical Report, TR37/04.

[22] P. Kulik (2000). Software Metrics: State of the Art. Retrieved 10/18/05 World Wide Web, http://www.klci.com

[23] A. Neely (1998). *Measuring Business Performance. Why, What and How?*. Profile Books Ltd, London.

[24] Software Productivity Center Inc. *SPC Metrics Recources, 8-step Metrics Program*. Retrieved 8/18/07 World Wide Web, http://www.spc.ca/resources/metrics/index.html

[25] R. van Solingen and E. Berghout (1999). *The Goal/Question/Metric Method: A Practical Guide for Quality Improvement of Software Development*. McGraw-Hill Publishing Company, London, UK.

[26] I. Becerra-Fernandez, A. Gonzalez and R. Sabherwal (2004). *Knowledge Management. Challenges, Solutions, and Technologies.* Prentice Hall, Upper Saddle River, NJ.

[27] M. R. Testa (2002). A Model for organization-based 360 degree leadership assessment. *Leadership & Organization Development Journal,* Vol. 23, no. 5, pp. 260-268.

[28] H. Harju and M. Koskela (2003). *Cost-effective reliability design and assessment of software. Part 2*. VTT – Research Notes 2193, Espoo, Finland.

[29] S. H. Kan (2002). *Metrics and Models in Software Quality Engineering, Second Edition.* Addison-Wesley Professional, Boston, MA.

[30] S. H. Kan, V. R. Basili and L. N. Shapiro (1994). An overview from the perspective of total quality management. *IBM System Journal*, Vol. 33, no. 1, pp. 4-19.

[31] J. S. Osmundson, J. B. Michael, M. J. Machniak and M. A. Grossman (2003). Quality management metrics for software development. *Information & Management,* Vol. 40, pp. 799-812.

[32] T. Conti (1993). *Building Total Quality: A guide for management.* Chapman & Hall, New York

[33] S. Zahran (1998). *Software Process Improvement: Practical Guidelines for Business* Success. Addison-Wesley Longman Ltd, Essex, UK.

[34] W. S. Humphrey (1989). *Managing the Software Process*. Addison-Wesley Longman Publishing Co, Boston, MA.

[35] N. Fenton, "Software Measurement: A Necessary Scientific Basis", *IEEE Transactions on Software Engineering,* Vol. 20, Issue 3, March 1994, pp. 199-206.

[36] N. Fenton, and N. Martin, "Software metrics: successes, failures and new directions", *Journal of Systems and Software*, Vol. 47, Issues 2-3, July 1999, pp. 149-157.

[37] FiSMA, Finnish Software Measurement Association. Retrieved 10/16/07 World Wide Web, http://www.fisma.fi/eng/index.htm

[38] TUT, Tampere University of Technology, Retrieved 9/22/07 World Wide Web, http://www.tut.fi/public

[39] UJ, University of Joensuu, Retrieved 9/22/07 World WideWeb, http://www.joensuu.fi/englishindex.html

[40] SME; Small and Medium size enterprise. Retrieved 5/23/07 World Wide Web, http://ec.europa.eu/enterprise/enterprise_policy/sme definition/index_en.htm

[41] J. Soini, V. Tenhunen and T. Mäkinen (2007). Managing and Processing Knowledge Transfer between Software Organizations: A Case Study. *Proceedings of the PICMET'07 Conference,* Portland, USA. pp. 1108-1113.

[42] J. Soini, T. Varkoi, V. Tenhunen, M. Tukiainen (2007). Empirical case study of software metrics using the SPICE framework. *Proceedings of the SPICE'07 Conference,* Seoul, Korea. pp. 22-28.

[43] ISO/IEC IS 15504-5:2006 (2006) Information Technology - Process Assessment - Part 5: An exemplar Process Assessment Model (SPICE).

[44] CMMI for Systems Engineering/Software Engineering, Version 1.02. CMU/SEI-2000-TR-029. Carnegie Mellon University, SEI.

[45] E. Rodenbach, F. van Latum and R. van Solingen (2000). SPI – A Guarantee for Success? – A Reality Story from Industry. *Proceedings of the PROFES 2000 Conference*. Eds. F. Bomarius and M. Oivo, Oulu, Finland, pp. 216-231.

[46] T. Dybå (2000). An Instrument for Measuring the Key Factors of Success in Software Process *Improvement. Empirical Software Engineering*, Vol. 5, no. 4. pp. 357-390.

[47] S. L. Pfleeger and H. D. Rombach (1994). Measurement Based Process Improvement. *IEEE Software,* July, pp. 9-11.

[48] J. Soini, V. Tenhunen and M. Tukiainen (2007). Software Metrics and Evaluation of Their Usefulness in Finnish Software Companies. *Proceedings of the IWSM Mensura 2007 Conference,* Palma de Mallorca, Spain. pp. 204-215.

[49] S. Badker (1991). *Through the Interface. A Human Activity Approach to User Interface Design*. Lawrence Erlbaum Associates. Inc., New Jersey.

[50] M. Huysman and D. de Wit (2002). *Knowledge sharing in practice*. Kluwer Academic Publishers, Dordrecht.

[51] R. Petty and J. Guthrie (2000). Intellectual Capital literature review – measurement reporting and management. *Journal of Intellectual Capital*. Vol. 1, no. 2, pp. 155-176.

[52] ISO 9000-3 (1997). *Guidelines for the application of ISO 9001:1994 to the development, supply,*

installation and maintenance of computer software. Internal Organization for Standardization.
[53] ISO/IEC 9126-3;2003 (2003). *Software engineering – Product quality – Part 2: External metrics*. Technical Report.
[54] ISO/IEC 9126-3;2003 (2003). *Software engineering – Product quality – Part 2: Internal metrics*. Technical Report.
[55] IEEE Standard 1061-1998 (1998). *IEEE Standard for Software Quality and Metrics Methodology*. IEEE-SA Standard Board.
[56] Software Productivity Research (SPR). Retrieved 11/12/07 World Wide Web, http://www.spr.com/products/default.shtm
[57] International Software Benchmarking Standards Group (ISBSG). Retrieved 11/12/07 World Wide Web, http://www.isbsg.org/isbsg.nsf/weben/Tools
[58] R. M. Cooke (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science*. Oxford University Press, NY, Oxford.
[59] M. Bertrand and S. Mullainathan (2001). Do People Mean What They Say? Implications for Subjective Survey Data. *The American Economic Review*, Vol. 91, no. 2, pp.67-72.

Appendix A

Spreadsheet-style interview form for collecting detailed metrics information [42].

	A	B	C	D	E	F	G	H
1	#	Name	Purpose	Type	Target	Application Domain	Formula	Values
2		*name of the meter*	*extensive description of the meter and its purpose*	*meter's type: Monitoring, Controlling, Predicting, Validating*	*what attributes of product, process and/or project are measured*	*in what processes, lifecycle phases etc. the meter is used*	*how the meter is calculated*	*values the meter produces and their interpretation: what's good, what's bad, what's preferred*

	I	J	K	L	M	N	O	P
1	Data 1	Data 2	Data 3	Data 4	Data Collection Rate	Primary Collectors	Secondary Collectors	Usage
	data used to calculate the meter	*data used to calculate the meter*	*data used to calculate the meter*	*data used to calculate the meter (add more columns if needed)*	*how often the data is collected (e.g. X times in a day/week/month/other, please specify)*	*who are primarily responsible for measuring or for collecting the data*	*who are secondarily responsible for measuring or for collecting the data*	*how the data is collected and the metric calculated*

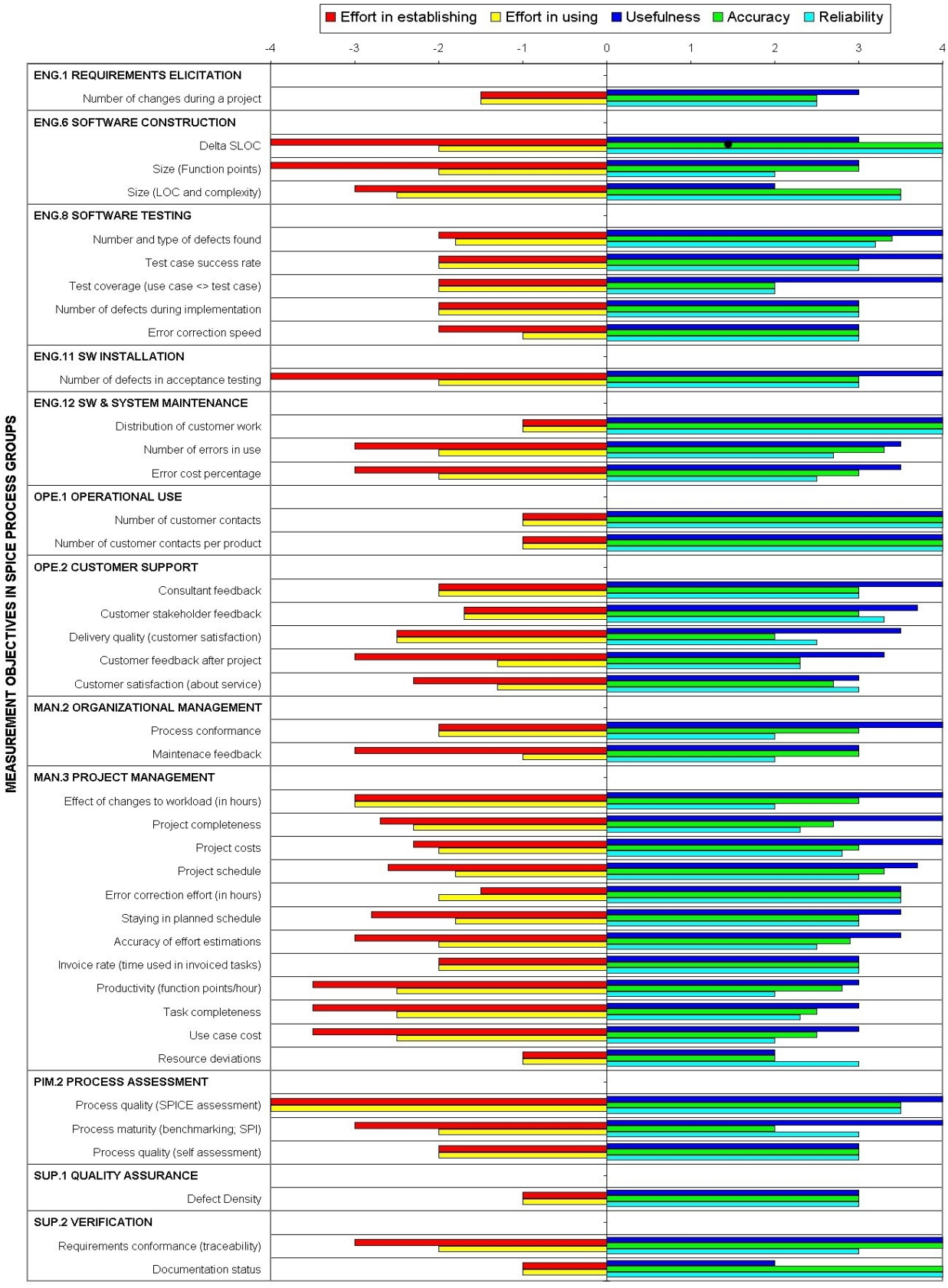
	Q	R	S	T	U
1	Examination Rate	Primary Beneficiaries	Secondary Beneficiaries	Workload in Establishing the Meter (1)	Workload in Establishing the Meter (2)
2	*how often are the meter's results looked at (e.g. X times in a day/week/month/other, please specify)*	*who are primarily using the meter's results*	*who are secondarily using the meter's results*	*how much resources are consumed when the meter is first introduced and established (e.g. person-hours, calendar time etc.)*	*estimation of the workload; scale: 1 = heavy, 2 = considerable, 3 = moderate, 4 = light*

	V	W	X	Y	Z	AA
1	Workload in Using the Meter (1)	Workload in Using the Meter (2)	Accuracy	Reliability	Risks	Usefulness
2	*how much resources are consumed when the meter is used (e.g. person-hours, calendar time etc.)*	*estimation of the workload; scale: 1 = heavy, 2 = considerable, 3 = moderate, 4 = light*	*estimated accuracy of the meter's results; scale: 1 = inaccurate, 2 = approximate, 3 = quite accurate, 4 = very accurate*	*estimated reliability and robustness of the meter's results; scale: 1 = unreliable, 2 = moderate, 3 = quite reliable, 4 = very reliable*	*risks and problems related to the meter's usage*	*general estimation of the meter usefulness; scale: 1 = useless, 2 = of limited use, 3 = quite useful, 4 = very useful*

	AB	AC	AD	AE	AF
1	Other Information	Source	References	Author	Web Links
2	*free-form notes, comments and other information (e.g. meter's relation to various process models; to different sizes and types of projects or organizations; etc.)*	*source of the meter, any of the following: Literature, Organizations, Standards*	*detailed sources of the meter: - name of the author and book/article/web site/etc. - name of the company or organization - name of the standard or model*	*name of the author who has written the information about the meter in this table*	*possible links to WWW sites with related information*

6 Appendix B

A summary of current measurement objects (for the SPICE framework viewpoint) with user evaluation based on five separate categories (effort requirements for establishing, effort requirements for use, reliability, accuracy and usefulness) [48]. The four -point scale for evaluating each category is given in Appendix A.



Study of Robust and Intelligent Surveillance in Visible and Multi-modal Framework

Praveen Kumar, Ankush Mittal and Padam Kumar
Department of Electronics and Computer Engineering,
Indian Institute of Technology, Roorkee,
India 247667
E-mail: praveen.kverma@gmail.com,{ankumfec,padamfec}@iitr.ernet.in

Keywords: video surveillance, object detection and tracking, data fusion, event detection

Received: December 21, 2006

This paper gives a review of current state of the art in the development of robust and intelligent surveillance systems, going beyond traditional vision based framework to more advanced multi-modal framework. The goal of automated surveillance system is to assist the human operator in scene analysis and event classification by automatically detecting the objects and analyzing their behavior using computer vision, pattern recognition and signal processing techniques. This review addresses several advancements made in these fields while bringing out the fact that realizing a practical end to end surveillance system still remains a difficult task due to several challenges faced in a real world scenario. With the advancement in sensor and computing technology, it is now economically and technically feasible to adopt multi-camera and multi-modal framework to meet the need of efficient surveillance system in wide range of security applications like security guard for communities and important buildings, traffic surveillance in cities and military applications. Therefore our review includes significant discussion on multi-modal data fusion approach for robust operation. Finally we conclude with discussion on possible future research directions.

Povzetek: Opisane so moderne robustne metode inteligentnega nadzora.

1 Introduction

Security of human lives and property has always been a major concern for civilization for several centuries. In modern civilization, the threats of theft, accidents, terrorists' attacks and riots are ever increasing. Due to the high amount of useful information that can be extracted from a video sequence, video surveillance has come up as an effective tool to forestall these security problems. The automated security market is growing at a constant and high rate that is expected to sustain for decades [1]. Video surveillance is one of the fastest growing sectors in the security market due to its wide range of potential applications, such as a intruder detection for shopping mall and important buildings [2], traffic surveillance in cities and detection of military targets[3], recognition of violent/dangerous behaviors (eg. in buildings, lifts) [4] etc. The projections of the compound annual growth rate of the video-surveillance market are about 23% over 2001-2011, to touch US$670.7 million and US$188.3 million in USA and Europe, respectively [5].

An automated surveillance system attempts to detect, recognize and track objects of interest from video obtained by cameras along with information from other sensors installed in the monitored area. The aim of an automated visual surveillance system is to obtain the description of what is happening in a monitored area and to automatically take appropriate action like alerting a human supervisor, based on the perceived description. Visual surveillance in dynamic scenes, especially for humans and vehicles, is currently one of the most active research topics in computer vision [6]. For at least two decades, the scientific community has been involved in experimenting with video surveillance data to improve image processing tasks by generating more accurate and robust algorithms in object detection and tracking [7,8], human activity recognition [9,10], database [11] and tracking performance evaluation tools [12].

The most desirable qualities of a video surveillance system are (a) *robust* operation in real world scenarios, characterized by sudden or gradual changes in the input statistics and (b) *intelligent* analysis of video to assist the operators in scene analysis and event classification. In the past several research works have been carried out in many fields of video surveillance using single vision camera and indeed significant results have been obtained. But mostly they are proven to work in a controlled environment and specific contexts. A typical example is of vehicle and traffic surveillance: systems for queue monitoring, accident detection, car plate recognition etc. In a recent survey on video surveillance and sensor networks research, Cucchiara [13] reports that there are still many unsolved problems in tracking in non ideal conditions, in cluttered and unknown environment, with variable and unfavorable luminance conditions, for

surveillance in indoor and outdoor spaces. Traditional approaches in dealing with these problems have focused on improving the robustness of background model and object segmentation techniques by extracting additional content from data (color, texture etc). However they have used only single modality such as visible spectrum or thermal infrared video. Visible and thermal infrared spectrums are intuitively complementary, since they capture information in emitted and reflected radiations, respectively. Thus alternative approach of integrating information from multiple video modalities has the potential to deal with such dynamically changing environment by leveraging the combined benefits whilst compensating for failures in individual modalities [14].

In addition other media streams like audio can improve analysis of visual data. For example, visual and ambient media capture two different aspects - scene and sound, respectively. In many cases where visual information is not sufficient for reliably discriminating between activities, there is often audio stimulus that is extremely important for a particular classification or anomaly detection task [15].

Automatic intelligent analysis of incoming video data on-line is required because firstly it is practically infeasible to manually supervise huge amount of video data (especially with multiple cameras) and secondly, off-line analysis completely precludes any possibility of taking immediate action in the likely happening of an abnormal event, particularly in critical applications. Several intelligent activity/ event detection methods are being proposed as the behavior patterns of real life scenario still remain challenge for the research community.

Therefore our emphasis in this paper is to discuss the existing (and proposed) techniques and provide summary of progress achieved in the direction of building robust and intelligent surveillance system. The paper's scope goes beyond traditional vision based framework to multi-modal framework. In several places, we briefly review some related concepts in automated surveillance system to put everything in proper context. For detailed discussion on studies in those related areas, reviews are available as follows: Background subtraction techniques [16], tracking of people and body parts [17], face recognition [18], gesture recognition [19], issues in automated visual surveillance [20], multimedia and sensor networks [13], distributed surveillance systems [21] and a detailed review of techniques in all the stages in the general framework of visual surveillance [6].

The rest of the paper is organized as follows. Section 2 gives an overview of automated visual surveillance, its evolution and practical issues. Section 3 discusses computer vision techniques in and beyond visual spectrum that have been developed for object detection and tracking. Section 4 reviews the work related to data fusion in multi modal framework (including visible, infrared and audio). Section 5 covers the activity recognition and behavior understanding approaches for event detection. Finally section 6 concludes the paper by summarizing the discussion and analyzing some possible future research directions.

2 Overview of automated visual surveillance system

The general framework of an automatic video surveillance system is shown in Figure1. Video cameras are connected to a video processing unit to extract high-level information identified with alert situation. This processing unit could be connected throughout a network to a control and visualization center that manages, for example, alerts. Another important component is a video database and retrieval tool where selected video segments, video objects, and related contents can be stored and inquired. In [6, 22], a good description of video object processing in surveillance framework is presented. The main video processing stages include background modeling, object segmentation, object tracking, behaviors and activity analysis. In multi camera scenario, fusion of information is needed, which can take place at any level of processing. Also these cameras may be of different modality like thermal infrared, near infrared, visible color camera etc so that multi spectral video of the same scene can be captured and the redundant information may be used to improve the robustness of the system against dynamic changes in environmental conditions.

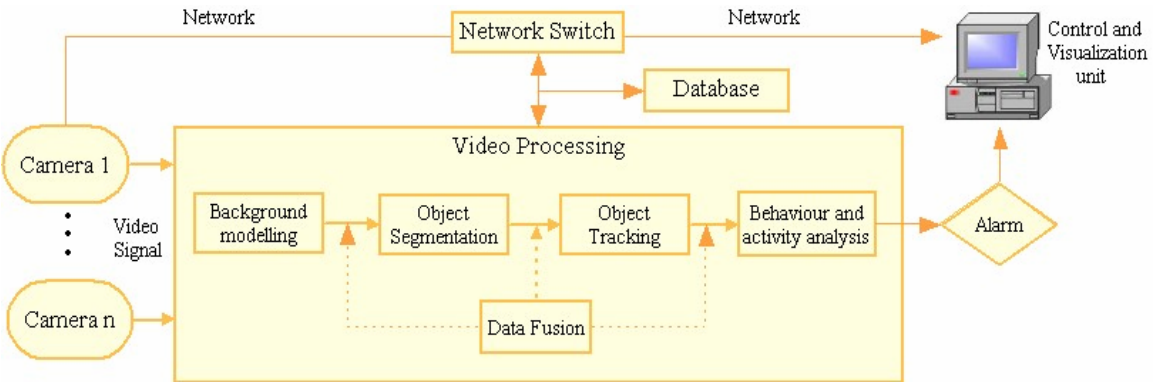


Figure1: General framework of automated visual surveillance system

2.1 Evolution of surveillance systems

"First generation" video-based surveillance systems started with analog CCTV systems, which consisted of a number of cameras connected to a set of monitors through automated switches. In [23], for example, integration of different CCTV systems to monitor transport systems is discussed. But the human supervision being expensive and ineffective due to widespread deployment of such systems, they are more or less used as a forensic tool to do investigation after the event has taken place. By combining computer vision technology with CCTV systems for automatic processing of images and signals, it becomes possible to proactively detect alarming events rather than passive recording. This led to the development of semi-automatic systems called "second generation" surveillance systems, which require a robust detection and tracking algorithm for behavioral analysis. For example, the real-time visual surveillance system W4 [7] employs a combination of shape analysis and tracking, and constructs models of people's appearances in order to detect and track groups of people as well as monitor their behaviors even in the presence of occlusion and in outdoor environments. Current research issues in such systems are mainly real time robust computer vision algorithms and automatic learning of scene variability and patterns of behaviors.

Third generation surveillance system is aimed towards the design of large distributed and heterogeneous (with fixed, PTZ, and active cameras) surveillance systems for wide area surveillance like monitoring movement of military vehicles on borders, surveillance of public transport etc. For example the Defense Advanced Research Projection Agency (DARPA) supported the Visual Surveillance and Monitoring (VSAM) project [24] in 1997, whose purpose was to develop automatic video understanding technologies that enable a single human operator to monitor behaviors over complex areas such as battlefields and civilian scenes. The usual design approach of these vision systems is to build a wide network of cooperative multiple cameras and sensors to enlarge the field of view. From an image processing point of view, they are based on the distribution of processing capacities over the network and the use of embedded signal processing devices to give the advantages of scalability and robustness potential of distributed systems. The main research problems involved in such systems are: integration of information obtained from different sensors, establishing signal correspondence in space and time, coordination and distribution of processing task and video communication etc.

Recently, the rapid emergence of wireless networks and proliferation of networked digital video cameras have favorably increased the opportunity for deploying large scale Distributed Video Surveillance (DVS) systems on top of existing IP-network infrastructure. Many commercial companies now offer IP-based surveillance solutions. For example companies like Sony and Intel have designed equipments like smart cameras; Cisco provides many networking devices for video surveillance. All this has led to the latest step in the evolution of video-surveillance systems i.e migration to digital IP-based surveillance and recently to wireless interconnection network. Figure 2 shows a general DVS network architecture, where there are several video sensors/cameras distributed over a wide area, with smaller groups under a local base station called Processing proxy server (PPS). A PPS collects video streams from many such video cameras through a wireless (mostly) or wired LAN or mesh network. These servers are equipped with computational power to perform necessary machine vision processing and data filtering to analyze the video stream and identify alert situations. These servers then transmit the video data to different users the backbone internet network.

2.2 Practical issues in real world scenario

Despite much advancement in the field, realizing practical an end-to-end video surveillance system in a real world scenario remains a difficult task due to the following issues:

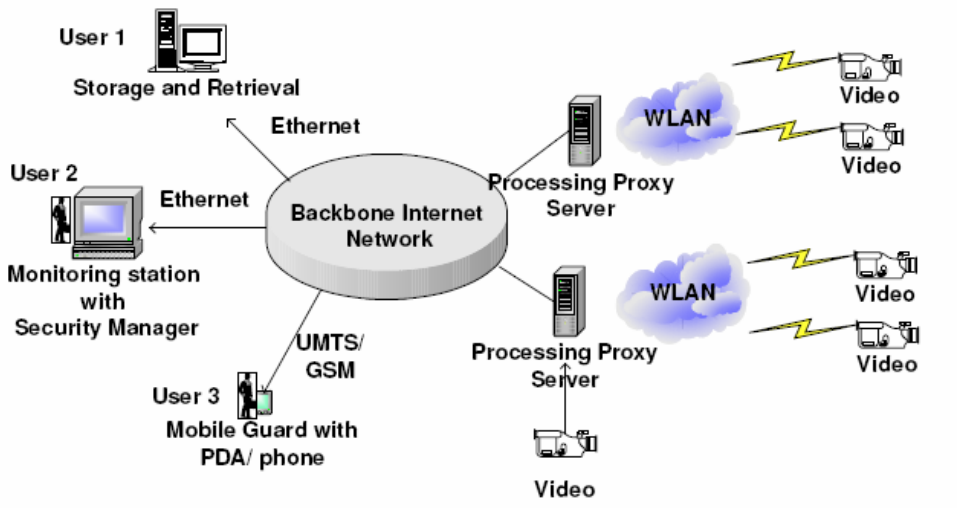


Figure 2. Distributed Video Surveillance Network Architecture

1. Robustness: Real world scenarios are characterized by sudden or gradual changes in the input statistics. A major challenge for real world object detection and tracking is the dynamic nature of real world conditions with respect to illumination, motion, visibility, weather change etc. As pointed out in [22], achieving robust algorithms is a challenge especially (a) under illumination variation due to weather conditions or lighting changes, for example, in outdoor scene, due to movement of clouds in sky and in an indoor scene, due to opening of doors or windows; (b) under view changes; (c) in case of multiple objects with partial or complete occlusion or deformation; (d) in the presence of articulated or non-rigid objects; (e) in case of shadow, reflections, and clutter; and (f) with video noise (e.g., Gaussian white). Figure 3 shows scenarios with (a) low light and illumination variation (b) video noise (c) boat among moving waves and (d) car object with moving background of vegetation. Significant research and advancement in solving these difficulties have been achieved but still the problem is unsolved in generic situation with dynamically varying environmental conditions and there is lack of generic multimodal framework to achieve system robustness by data fusion.

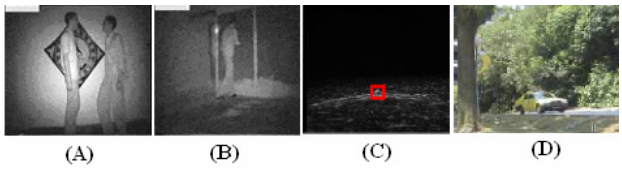


Figure 3. Some examples of complex real world situation for object detection

2. Intelligent: With the advances in sensor technology, surveillance cameras and sound recording systems are already available in banks, hotels, stores and highways, shopping centers and the captured video data are monitored by security guards and stored in archives for forensic evaluation. In a typical system, a security guard watches 16 video channels at the same time and may miss many important events. There is a need of intelligent, (semi-) automated video analysis paradigms to assist the operators in scene analysis and event classification. Event detection is a key component to provide timely warnings to alert security personnel. It deals with mapping motion patterns to semantics (e.g., benign and suspicious events). However detecting semantic events from low-level video features is major challenge in real word situation due to unlimited possibilities of motion patterns and behaviors leading to well known semantic gap issue. Furthermore, suspicious motion events in surveillance videos happen rather infrequently and the limited amount of training data poses additional difficulties in detecting these so-called rare events [25].

3. Real timeliness: A useful processing algorithm for surveillance systems should be real time, i.e., output information's, such as events, as they occur in the real scene [22]. Requirement of accuracy and robustness result in computational intensive and complex design of algorithms which makes real time implementation of system a difficult task.

4. Cost effective: For feasible deployment in a wide variety of real world surveillance applications ranging from indoor intrusion detection to outdoor surveillance of important buildings etc, a cost effective framework is required.

3 Computer vision techniques for visual surveillance tasks

This section summarizes the research that addresses the basic computer vision problems in video surveillance like object detection and tracking. These modules constitute the low level building block necessary for any surveillance system and we briefly outline the most popular techniques used in these modules. We also present the advances made in computer vision techniques both in and beyond the visible spectrum (thermal infrared etc.) to give motivation for the discussion on data fusion in the next section.

3.1 Object detection

Nearly every visual surveillance system starts with object detection. Object detection aims at segmenting regions corresponding to moving objects such as vehicles and humans from the rest of an image. Detecting moving regions provides a focus of attention for later processes such as tracking and behavior analysis because only these regions need be considered in the later processes. There are two main conventional approaches to object detection: 'temporal difference' and 'background subtraction'. The first approach consists in the subtraction of two consecutive frames followed by thresholding. The second technique is based on the subtraction of a background or reference model and the current image followed by a labeling process. After applying one of these approaches, morphological operations are typically applied to reduce the noise of the image difference (See figure 4 and 5). The temporal difference technique is very adaptive to changes in dynamic environment and another advantage is that it does not make assumptions about the scene. However, it can be problematic that only motion at edges is visible for homogeneous objects. On the other hand, background subtraction has better performance extracting object information but it is sensitive to dynamic changes in the environment.

Figure 4: Example of object detection using temporal differencing technique

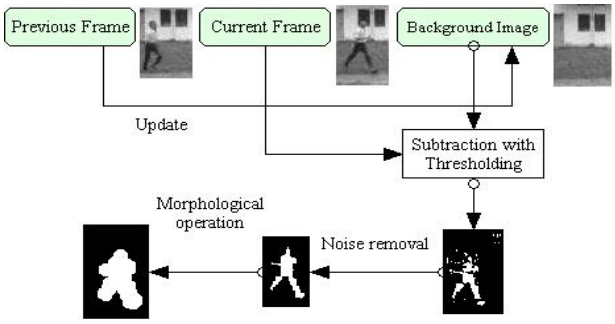


Figure 5: Object Detection using background subtraction technique

Background modeling assumes that the video scene is composed of a relatively static model of the background, which becomes partially occluded by objects that enter the scene. These objects (usually people or vehicles) are assumed to have features that differ significantly from those of the background model (their color or edge features, for example). The terms *foreground* and *background* are not scientifically defined however and thus their meaning may vary across applications. For example, a moving car should usually be considered as a foreground object but when it parks and remains still for a long period of time, it is expected to become background. Also, not all moving objects can be considered foreground. The simplest approach is to record an image when no objects are present and use this image as the background model. However, continuous updating of the model is required to make the foreground extraction more robust to the gradual changes in lighting and movement of static objects that are to be expected in outdoor scenes. Unfavorable factors, such as illumination variance, shadows and shaking branches, bring many difficulties to the acquirement and updating of background model. Background modeling is a very active research area and several techniques have been proposed to deal with various problems. A good overview of the most frequently cited background modeling algorithms is given in [16]. A comparison between various background modeling algorithms is given in [26], as well as a discussion on the general principles of background maintenance systems.

A typical approach for modeling background in outdoor conditions is using Gaussian model that models the intensity of each pixel with a single Gaussian distribution [27] or with more than one Gaussian distribution. The algorithm described in [28] models each pixel as a sum of K Gaussian distributions in RGB space ($1 \leq K \leq 5$). Each pixel's background model is updated continuously, using online estimation of the parameters. This model is well suited to cater for pixels whose background model has a multimodal distribution, such as vegetation or water. The model is unable to distinguish between foreground objects and shadows, however, and also is quite slow to initialize. The algorithm used in the $W4$ [6] system works on monochrome video and marks a pixel as foreground if:

$$|M - It| > D \text{ or } |N - It| > D \text{ (1)} \quad (1)$$

where the (per pixel) parameters M, N, and D represent the minimum, maximum, and largest inter-frame absolute difference observed in the training frames. It also detects when its background model is invalid by detecting when 80% of the image appears as foreground. To rectify this, is re-enters a training mode to correct the background model. However reliable background modeling is difficult to achieve in certain scenarios. For example, in a crowded room with many people, the background may only ever be partially visible. Another problematic scenario is in a scene with low levels of lighting, such as a night-time scene with only street lighting. The movement (or apparent movement) of background objects is problematic too. Examples of this include moving trees and vegetation, flickering computer or TV screens, flags or banners blowing in the wind, etc.

Apart from common approached discussed above, techniques based on optic flow is useful in motion segmentation where a motion vector is assigned to every pixel of the image by comparison of successive frames. Optical-flow-based methods can be used to detect independently moving objects even in the presence of camera motion. However, most flow computation methods are computationally complex and very sensitive to noise, and cannot be applied to video streams in real time without specialized hardware. More detailed discussion of optical flow can be found in Barron's work [29].

3.2 Object tracking

Once objects have been detected, the next logical step is to track these detected objects. Tracking has a number of benefits. Firstly, the detection phase is quite computationally expensive, so by using tracking, the detection step does not need to be computed for each frame. Secondly, tracking adds temporal consistency to sequence analysis because otherwise, objects may appear and disappear in consecutive frames due to detection failure. Also, tracking can incorporate validity checking to remove false positives from the detection phase. Thirdly, if tracking multiple objects, detection of occlusion is made easier, as we expect occlusion when two or more tracked object move past each other (as shown in figure 6). Object motion can be perceived as a result of either camera motion with a static object, object motion with static camera, or both object and camera moving. Tracking techniques can be divided into two main approaches: 2-D models with or without explicit shape models and 3-D models. For example, in [30] the 3-D geometrical models of a car, a van and a lorry are used to track vehicles on a highway. The model-based approach uses explicit a priori geometrical knowledge of the objects to follow, which in surveillance applications are usually people, vehicles or both. In [6], author use a combination of shape analysis and along with 2D Cardboard Model for representing and tracking the different body parts. Along with second order predictive motion models of the body and its parts, they used

Cardboard Model to predict the positions of the individual body parts from frame to frame.

A common tracking method is to use a filtering mechanism to predict each movement of the recognized object. The filter most commonly used in surveillance systems is the Kalman filter [31]. Fitting bounding boxes or ellipses, which are commonly called 'blobs', to image regions of maximum probability is another tracking approach based on statistical models. In [27] the author models and tracks different parts of a human body using blobs, which are described in statistical terms by a spatial and color Gaussian distribution. In some situations of interest the assumptions made to apply linear or Gaussian filters do not hold, and then nonlinear Bayesian filters, such as extended Kalman filters (EKF) or particle filters have been proposed. A good tutorial on non linear tracking using particle filter is given in [32] where the author illustrates that in highly non-linear environments particle filters give better performance than EKF. A particle filter is a numerical method, which weights (or 'particle') a representation of posterior probability densities by resampling a set of random samples associated with a weight and computing the estimate probabilities based on these weights. Then, the critical design decision using particle filters relies on the choice of importance (the initial weight) of the density function. Appearance models [33] are another way to represent objects. It consists of an observation model (usually an image) of the tracked object, along with some statistical properties (such as the pixel variances).

Figure 6: Object Tracking during and after Occlusion

3.3 Computer vision beyond the visible spectrum

Traditionally, the majority of the computer vision community has been involved implicitly or explicitly with the development of algorithms associated with sensors that operate in the visible band of the electromagnetic spectrum [34]. In the past, imaging sensors beyond visible spectrum have been limited to special applications like remote sensing and vision based military applications, because of their high cost. Recently with the advances in sensor technologies, the cost of near and mid-infrared sensors has dropped dramatically, making it feasible for their use in more common applications like automatic video-based security and surveillance systems to enhance their capabilities.

Lin, in his 2001 technical report explores the extension of visible band computer vision techniques to infrared as well as conducts a good review of infrared imaging research [35]. Recent literature on the exploitation of near-infrared information to track humans generally deals only with the face of observed people[36] and a few are concerned with the whole body [37,38] but these approach rely on the highly limiting assumption that the person region always has a much brighter (hotter) appearance than the background. This assumption does not hold in various weather conditions and during all time. To tackle this, the author in [39] proposes a novel contour based background subtraction strategy to detect people in thermal imagery, which is robust across a wide range of environmental conditions. First of all, a standard background-subtraction technique is used to identify local region-of interest (ROI), each containing the person and surrounding thermal halo. The foreground and background gradient information within each region are then combined into a contour saliency map (highlighting the person boundary). Using a watershed-based algorithm, the gradients are thinned and thresholded into contour fragments. The remaining watershed lines are used as a guide for an A* search algorithm to connect any contour gaps. Finally, the closed contours are flood-filled to make silhouettes.

The use of infrared in pedestrian detection to reduce night time accidents is investigated in [38]. In [40], the author investigates human repetitive activity properties using thermal imagery. They employ a spatio-temporal representation, Gait Energy Image (GEI), which represent human motion sequence in a single image while preserving some temporal information. However they have developed the method for only simple activities which are repetitive in nature (like walking, running etc).

4 Data fusion

A surveillance task using multiple modalities can be divided into two major phases: *data fusion* and *event recognition*. The data-fusion phase integrates multi-source spatio-temporal data to detect and extract motion trajectories from video sources. The event-recognition phase deals with classifying the events as to relevance for the search. This section discusses the data fusion part and the event-recognition task is discussed in the next section.

Data fusion is the process of combining data from multiple sources such that the resulting entity or decision is in some sense better than that provided by any of the individual sources. Most of the existing surveillance systems have used only one media (i.e. normal video), and therefore they do not capture different aspects of the environment. Multiple media are useful because each media captures different aspect of the environment. For example sensing environmental sound can provide reliable clue for detecting insecure events in many cases. Infrared is more informative in dark environment, especially at night. Visible and thermal infrared spectrums are intuitively complementary, since they capture information in emitted and reflected radiations, respectively. Thus combining them can be advantageous in many scenarios, especially when one modality perform poorly in detecting objects. For example visible analysis has an obvious limitation of daytime operation only and completely fails in total darkness. Additionally foggy weather condition, sudden lighting changes,

shadows and color camouflage, often cause poor segmentation of actual objects and much false positive detection. Thermal infrared video is almost completely immune to lighting changes, and thus it is very robust to the above mentioned problems. However, infrared video has its unique inherent challenges due to high noise, and "Halo effect" produced by some infrared sensors, which appears as a dark or bright halo surrounding very hot or cold objects respectively. Further if people are wearing insulated clothing or infrared camera performs rapid automatic gain then it will cause foreground detection to incorrectly classify pixels. See figure 7 for illustration in two different situations. Thus by data fusion approach, it is possible to improve robustness of the system in dynamic real world conditions.

4.1 Fusion of visible and infrared

Depending on the application and fusion method, research in the fusion of visible and infrared imagery can be classified in two broad categories. *Image based Representational fusion* and *Video based Analytical fusion*. In *Image based Representational fusion,* the goal is to obtain best representation of the data in a single image for improved visual perception by combining multiple images to create a single fused image that somehow represents the information content of the input images. This type of fusion is generally used for remote sensing and military applications. Depending on the synergy of the information inherent in the data, it may be possible to reduce noise, to extend the field of view beyond that of any single image, to restore high frequency content, and even to increase spatial resolution [41]. Image fusion techniques have had a long history in vision. Gradient-based techniques examining gradients at multiple resolutions [42] and several region-based multi resolution algorithms have been proposed such as the pyramid approaches of [43, 44] and the wavelet-based approach of [45].

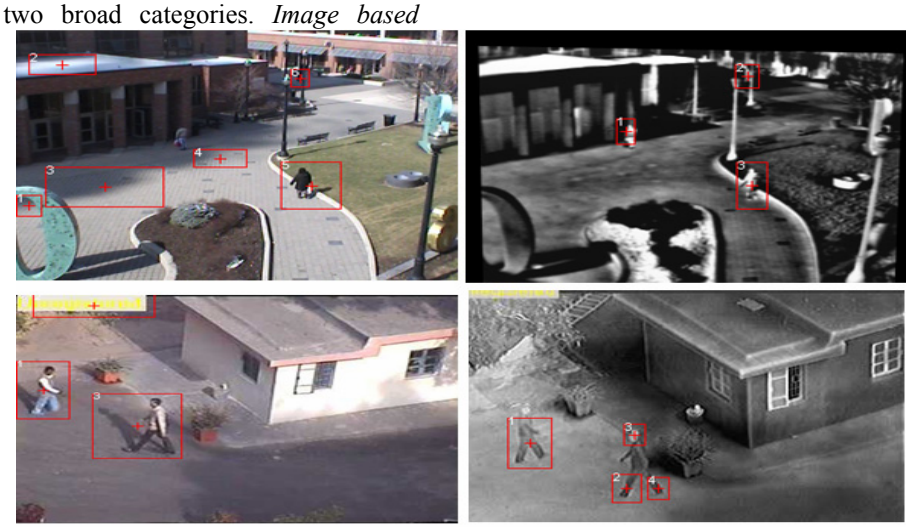


Figure 7: Visible and corresponding thermal infrared in a) Variable illumination due to cloud movement causing false detection in visible (top) b) Incorrect detection due to shadows in visible and thermally insulated clothing in infrared (bottom)

Video based Analytical Fusion, on the other hand, aims to extract knowledge by using all sources of data for better analysis, and not merely to represent the data in another way. This type of fusion methodology is required to enhance the capabilities of automatic video-based detection and tracking system for surveillance purpose. Although image fusion has received considerable attention in the past, research in the fusion of video modalities for automatic analysis, or analytical fusion is very recent. Some recent works have addressed the tracking of humans and vehicles with multiple sensors [46, 47] but issues that are involve in fusing multiple modalities for robust detection and tracking is very sparse. In [48], the fusion of thermal infrared with visible spectrum video, in the context of surveillance and security, is done at the object level. Detection and tracking of blobs (regions) are performed separately in the visible and thermal modality. An object is made up of one of more blobs, which are inherited or removed as time passes. Correspondences are obtained between objects in each modality, forming a master-slave relationship, so that the *master* (the object with the better detection or *confidence*) assists the tracking of the slave in the other modality. Their system uses many heuristics and there also seems to be many parameters to set empirically.

Davis et al. [49] propose a new contour-based background-subtraction technique using thermal and visible imagery for persistent object detection in urban settings. They perform statistical background subtraction in the thermal domain to identify the initial regions-of-interest. Color and intensity information are used within these areas to obtain the corresponding regions of-interest in the visible domain. Within each image region (thermal and visible treated independently), the input and background gradient information are combined as to highlight only the boundaries of the foreground object. The boundaries are then thinned and thresholded to form binary contour fragments. Contour fragments belonging to corresponding regions in the thermal and visible domains are then fused using the combined input gradient information from both sensors. An A* search

algorithm constrained to a local watershed segmentation is then used to complete and close any contour fragments. Finally, the contours are flood-filled to make silhouettes.

In a very recent work [50], the authors use thermal infrared video with standard CCTV video for object segmentation and retrieval in surveillance video. They segment object using separate background modeling in each modality and dynamic mutual fusion based thresholding. Transferable Belief Model is used to combine the sources of information for validating the tracking of objects. Extracted objects are subsequently tracked using adaptive thermo-visual appearance. However they don't take into account the reliability of each source in the fusion process. In [51], an intelligent fusion approach using Fuzzy logic and Kalman filtering technique is discussed to track objects and obtain fused estimate according to the reliability of the sensors. Appropriate measurement parameters are identified to determine the measurement accuracy of each sensor. A comparison of multiple fusion schemes for appearance based tracking of objects using thermal infrared and visible modalities is done in [52] for different objects, such as people, faces, bicycles and vehicles.

4.2 Data fusion methods

For visual surveillance using multiple cameras, issues such as camera calibration and registration, establishing correspondences between the objects in different image sequences taken by different cameras, target tracking and data fusion need to be addressed. The success of information fusion depends on how well data are represented, how reliable and adequate the model of data uncertainty used and how accurate and applicable prior knowledge is. Three commonly used fusion approaches are probabilistic methods (Bayesian inference), fuzzy logic method and belief models (Dempster-Shafer model and Transferable Belief model). The Bayesian inference method quantitatively computes the probability that an observation can be attributed to a given assumed hypothesis but lacks in ability to handle mutually exclusive hypotheses and general uncertainty [53]. Fuzzy logic methods accommodate imprecise states and variables. It provides tools to deal with observations that is not easily separated into discrete segments and is difficult to model with conventional mathematical or rule-based schemes [54]. The Belief theory generalizes Bayesian theory to relax the Bayesian method's restriction on mutually exclusive hypotheses, so that it is able to assign evidence to 'propositions', i.e. unions of hypotheses. Dempster-Shafer model makes a closed world assumption, so it assigns a belief of empty set to zero. The reasoning model assumes completeness of the frame of discernment meaning that the frame includes all hypotheses. But it can very well happen that some hypotheses, because of measurements are excluded from frame of discernment or unknown. In this way meaning of empty set is changed corresponding not only for impossibilities but also for unknown possibilities [55]. This kind of approach is called open world assumption which is considered in another belief model called Transferable Belief Model (TBM) [56]. TBM offers the flexibility to model closed world or open world assumption.

In [57], Bayesian probability theory is used to fuse the tracking information available from a suite of cues to track a person in 3D space. In [58], the authors uses TBM framework to solve the problem of data association in a multi target detection problem. It uses the basic belief mass $m(\Theta)$ as a measure of conflict and the sensors are clustered so that the conflict is minimized. But they tackle only partial problem of assessing how many objects are present and observed by the sensors. In [59], the author use TBM and Kalman filter for data fusion in object recognition system that analyses simulated FLIR and LADAR data to recognize and track aircraft. They demonstrated the results on an air to air missile based simulation system. In [60], the author proposes a hybrid multi-sensor data fusion architecture using Kalman filtering and fuzzy logic techniques. They feed the measurement coming from each sensor to separate fuzzy–adaptive kalman filters (FKF), working in parallel. The adaptation in each FKF is in the sense of adaptively adjusting the measurement noise coar1ance matrix R employing a fuzzy inference system (FIS) based on a covariance matching technique. Another FIS, which they call as fuzzy logic observer (FLO) monitors the performance of each FKF. Based on the value of a variable called Degree of Matching *(D0M)* and the matrix R coming from each FKF, the FLO assigns a degree of confidence, a number on the interval (0, 1], to each one of the FKFs output. The degree of confidence indicates to what level each FKF output reflects the true value of the measurement. Finally, a defuzzificator obtains the fused estimated measurement based on the confidence values. They demonstrated the result theoretically, by taking example of four noisy inputs.

4.3 Reliability of sensor

In the fusion process, different sources may have different reliability and it is essential to account for this fact to avoid decreasing in performance of fusion results. The fused estimate should be more biased by accurate measurements and almost unaffected by inaccurate or malfunctioning ones. Therefore for fusing data collected from different sensors requires the determination of measurements' accuracy so that they can be fused in a weighted manner. The most natural way to deal with this problem is to establish reliability of the beliefs computed within the framework of the model selected. For example [61] discusses a method for assessing the reliability of a sensor in a classification problem within the TBM framework. In [62], the authors propose a multi-sensor data fusion method for video surveillance, and demonstrated the results by using optical and infrared sensors. The measurements coming from different sensors were weighted by adjusting measurement error covariance matrix used by the fusion filter. To estimate the reliability of the sensor they defined a metric called Appearance Ratio (AR), whose value is proportional to

the strength of the segmented blobs from each sensor. The ARs are compared to determine which sensors are more informative and therefore selected to perform a specific video surveillance task. In [63], the authors discuss the principal concepts and strategies of incorporating reliability into classical fusion operators and provide good literature survey on main approaches used in fusion literatures to estimate reliability of sensor.

4.4 Audio and video information fusion

Enhancing visual data with audio streams can serve manifold purpose like speaker tracking, environment sound recognition for event recognition in surveillance application etc. Environmental sound like that of breaking of glass, dog's barking, screaming of a person, fire alarm, gun firing and similar kind of sounds, if detected and recognized correctly, can give a reasonable degree of confidence in making a decision about 'secure' vs. 'insecure' state [64]. Multimedia researchers have often used early fusion strategy to perform the audio-visual fusion for various problems including speech processing [65] and recognition [66], speaker localization [67] and tracking [68, 69], and monologue detection [70]. In [68], the authors present a method that fuses 2-D object shape and audio information via importance filters. They used audio information to generate an importance sampling function, which guides the random search process of particle filter towards regions of the configuration space likely to contain the true configuration (a speaker). A recent work in [71], describes a process to assimilate data from coarse and medium grain sensors, namely video and audio, and a probabilistic framework to discriminate concurring and contradictory evidences. The authors enlarge the concept to information fusion with the definition of *information assimilation*: this process includes not only the real-time information fusion but also the integration with the past experience, represented by the surveillance information stored in the system.

Research in the field of environmental sound recognition is sparse. The majority of auditory research is centered on the identification and recognition of speech signals. Those systems that do exist, work on a very specialized domain like in [72], a system named *AutoAlert* is presented for automated detection of incidents using HMM, and Canonical Variates Analysis (CVA) to analyze both short-term and time-varying signals that characterize incidents. Cowling, in [73] provides more detailed literature survey, and it also investigates few existing techniques used for sound recognition in speech and music. He then presents a comparison on the accuracy of these techniques, when employed for the problem of non-speech environmental sound classification for autonomous surveillance.

4.5 Other sensors/modalities

There are proximity sensors (like ultrasound devices, lasers scanners etc.) which detect objects without physical contact. Most proximity sensors emit an electromagnetic field or beam and look for changes in the field. Different targets demand different sensors. For example, a capacitive or photoelectric sensor might be suitable for a plastic target; an inductive sensor requires a metal target [74]. In [75], the authors integrate Laser Doppler Vibrometer (LDV) and IR video for remote multimodal surveillance. Their work mainly caters to remote area surveillance and their main focus was to study the application of LDV for remote voice detection, while IR imaging was used for target selection and localization. Few object tracking and visual servoing system for the visually impaired such as the *GuideCane* [76] and the *NavBelt* [77], use ultrasound or laser rangefinders to detect obstacles.

Gated imaging is another useful system for highly unfavorable visual conditions like underwater surveillance etc. Time gating is a temporal example of image formation whereby a light source is time pulse projected toward a target and the detector is time gated to accept image-forming illumination from a specific range [78]. LIDAR systems [79] time gate the receiver aperture to eliminate relatively intense backscatter originating from the water while allowing the return from the target to be detected. In [78] time gating is employed for using spatially and temporally varying coherent illumination for undersea object detection.

5 Event detection

5.1 Human activity recognition

Computer analysis of human actions is gaining increasing interests, especially in video surveillance arenas where people identification and activity recognition are important. Using two important metrics: preciseness of the analysis outcome and the required video resolution to achieve the desired outcome, human identification and activity recognition can be classified into three categories. At one extreme, which is often characterized by high video resolution and a small amount of scene clutter, high fidelity outcome is achievable. Many techniques in face, gesture, and gait recognitions fall in this category, which aim to identify individuals against a pre-established database.

At the other extreme, which is characterized by low video resolution and potentially significant scene clutter, it is often not possible to achieve highly discriminative outcome. Instead, the goal is often to detect the presence, and identify the movement and interaction of people through "blob" tracking [6, 24]. The VSAM system [24] tracked the human body as a whole blob. They use a hybrid algorithm by combining adaptive background subtraction with a three-frame differencing technique to detect moving objects, and use the Kalman filter to track the moving objects over time. A neural network classifier is trained to recognize four classes: single person, group of persons, vehicles, and clutter. They also use linear discriminant analysis to further provide a finer distinction between vehicle types and colors. The VSAM system is very successful at tracking humans and cars, and at discriminating between vehicle types. But it did not put

much emphasis on activity recognition; only gait analysis and simple human-vehicle activity recognition are handled.

In the middle of the spectrum, it is possible to refine the "blob" representation of a person through hierarchical, articulated models. For example, [80] describes an approach which attempts to recognize more generic activities and movement of body parts using MHI (motion-history images) to record both the segmentation result and the temporal motion information. The MHI is a *single* image composed of superimposing a sequence of segmented moving objects weighed by time. The most recent foreground pixels are assigned the brightest color while past foreground pixels are progressively dimmed. This allows the summarization of information on both the spatial coverage and the temporal ordering of the coverage of an activity. See figure 8 for illustration with few examples. The MHI does not use any structure to model human. A vector of seven moment values is computed for each MHI. Activities are recognized by finding the best match of the moment vectors between the query MHI and the training patterns. Other approaches allow main body parts, such as head, arms, torso, and legs, to be individually identified to specify the activities more precisely. A very detailed review of activity recognition approaches in these categories can be found in [8].

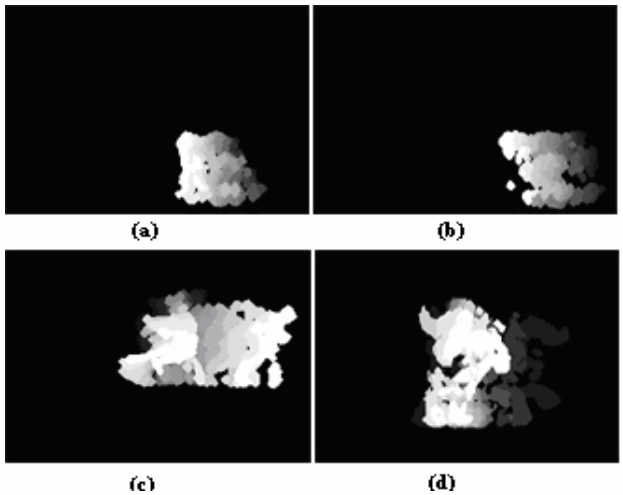


Figure 8. MHI images of person (a) walking (b) Running, (c) picking an object and (d) fighting

5.2 Semantic information extraction for behavior understanding

Understanding of behaviors may simply be thought as the classification of time varying feature data, i.e., to analyze the video to extract some feature vectors and to classify this time-varying feature data. During the recognition phase, extracted unknown test feature vector set is compared to a group of labeled reference feature vector sets representing typical human actions.

Feature extraction process is very important to achieve good results. It is impossible to train and perform classification using the currently available classification engines. The size of the feature vector should be as small as possible to have computational efficiency. At the same time it should represent each action very accurately. For example, the center of mass of the tracked object in image frame of the video can be used as a feature vector in a security application in which moving persons entering or leaving a building. In this simple problem, the "vocabulary" consists of a person leaving the building and a person entering a building and the center of mass information consisting of the horizontal and vertical coordinates in an image of the video may be good feature vector for this problem. On the other hand, detecting an 'assault' case where a person falls on the ground and other runs will require other additional parameters in the feature set. For example, so-called snaxels of an active snake contour of the human body can be added to the feature vector to distinguish a fallen person from a person standing. Compactness of the contour boundary, the speed of the center of the mass etc can be also used to distinguish the normal action of walking and the abnormal action of a falling and running. As a rule of thumb, model parameters are selected as entries of the feature vector in a model based tracking approach. Since a video consists of sequence of images a sequence of feature vectors are obtained to characterize the motion of a person(s).

5.3 Pattern analysis and classification methods

Several generative and discriminative models have been proposed for modeling and classifying activity patterns. Some of the most widely used ones are Dynamic time Warping (DTW), Hidden Markov Models (HMM), Time Delay Neural Network (TDNN), and Finite State Machine (FSM) network.

a) Dynamic time warping: DTW is a template based dynamic programming matching technique widely used in the algorithms for speech recognition. It has the advantage of conceptual simplicity and robust performance, and has been used recently in the matching of human movement patterns [81]. For instance, Bobick *et al.* [82] use DTW to match a test sequence to a deterministic sequence of states to recognize human gestures. Even if the time scale between a test sequence and a reference sequence is inconsistent, DTW can still successfully establish matching as long as the time ordering constraints hold.

b) Hidden Markov Models: HMMs are stochastic finite state machines [83]. In the context of human motion analysis, a finite state Markov model is assigned for each possible scenario and its parameters are trained with feature vectors of this typical human action. The training process is an off-line iterative algorithm called Baum-Welch algorithm [84]. Here , the number of states of a HMM must be specified, and the corresponding state transition and output probabilities are optimized in order that the generated symbols can correspond to the observed image features of the examples within a specific movement class. During the classification or recognition phase, the test feature vector set is applied to

all of the Markov models and output probabilities are computed. The Markov model producing the highest probability is determined and the corresponding human action scenario is selected as the result. HMMs generally outperform DTW for undivided time series data, and are therefore extensively applied to behavior understanding. In [85], authors describe an activity recognition process for visual surveillance of wide areas and experimented with image sequences acquired from an archaeological site with actors perform both legal and illegal actions. The activity recognition process is performed in three steps: first of all the binary shape of moving people are segmented, then the human body posture is estimated frame by frame and finally, for each activity to be recognized, a temporal model of the detected postures is generated by Discrete Hidden Markov Models

c) Finite state machine: The most important feature of a FSM is its state-transition function. The states are used to decide which reference sequence matches with the test sequence. However it requires hand crafted heuristic rules based on context knowledge. For example in [86], the authors propose a framework for unsupervised learning of usual activity patterns and detection of unusual activities based on a model of multi-layered finite state machines. They considered two different approaches for different scenario. First approach is unsupervised learning of usual activity patterns and detection of unusual activities (which are not recognized as normal). Other approach is to explicitly program the FSM or training using supervised learning for recognition of situation specific activities like unattended baggage detection.

d) *Time-delay neural network (TDNN):* TDNN is also an interesting approach to analyzing time-varying data. In TDNN, delay units are added to a general static network, and some of the preceding values in a time-varying sequence are used to predict the next value. As larger data sets become available, more emphasis is being placed on neural networks for representing temporal information. TDNN has been successfully applied to hand gesture recognition [87] and lip-reading [88].

Other related schemes including Dynamic Bayesian Network (DBN) and the Support Vector Machines (SVM) are also now being actively used in pattern analysis and classification problems.

Considering the limited training data for unusual events and that the distinction between two unusual events can be as large as those between unusual events and usual events, it is not feasible to train a general model for the unusual events. Therefore alternative approach that some people have taken is to train a model for usual events and events that deviate significantly from the usual event model are considered unusual. For example a simple ATM surveillance scenario is shown in figure 9, where FSM can be a useful way to discriminate between normal and abnormal patterns. The first row shows frames corresponding to normal transaction. Second and third row shows event corresponding to vandalism and robbery.

Figure 9. Sampleshots for events in a Bank ATM Surveillance

Figure 10 shows a possible FSM for activities which will be considered normal if the transitions terminated at the exit node. If exit is made by another node due to any deviant pattern, the FSM will flag an abnormal event.

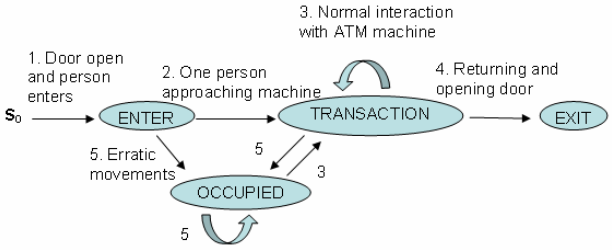


Figure 10. A possible FSM for Bank ATM Surveillance

Boiman and Irani [89] proposed to model the set of usual events as an ensemble of spatial-temporal image patches, and detect irregularity in a test video by evaluating the similarity between the test ensemble with the training ensemble. Zhang *et al.* [90] used a semi-supervised approach to train models for both usual and unusual events. They start from an ergodic hidden Markov model (HMM) for usual events. If a test event does not fit the model, they classify it as unusual and branch the usual event model to refit the usual event. This approach has a disadvantage that it may give high false positive alarm because in a practical scenario, even during the normal course of activity, unusual deviations are very likely to happen without potential threat.

6 Conclusions and future research developments

Event though significant progress have been made in computer vision and other areas, there are still major technical challenges to be overcome before the dream of reliable automated surveillance is realized. These technical challenges are compounded by practical considerations such as robustness to unfavorable weather and lighting conditions, intelligent video processing for event detection and efficiency in terms of real time

operation and cost. Most surveillance systems operate using single modality and lack robustness because they are limited to a particular situation. Different media source like audio, video and thermal infrared gives complementary/supplementary surveillance information of the environment. The literature survey on multi-modal data fusion shows that effective fusion of the information coming from different media streams can give robustness to real world object detection, tracking and event detection. Simultaneous fusion of thermal infrared and visible spectrum can give following advantage:

- Improved robustness against camouflage as foreground object are less likely to be of similar *color and temp* to the background
- Providing features that can be used for classification or retrieval of objects in large surveillance video archives
- Extracting a signature of an object in each modality, which indicates how useful each modality is in tracking that object

However, the key challenge for future research is to: 1) develop analysis techniques to automatically determine the reliability of data source and 2) develop a suitable fusion methodology that would intelligently utilize the information provided by these two modalities to get the best possible output. In [91], these challenges are addressed by employing Transferable Belief Model (TBM) and Kalman filter. TBM is used to determine the validity of a foreground region, detected by each source, for tracking. Kalman filter is used for the dual purpose of tracking the objects over time and fusing the measurements of the positions of the target obtained from different sensors, according to their reliability.

One of the main objectives of visual surveillance is to analyze and interpret individual behaviors and interactions between objects for event detection. Recently, related research has still focused on some basic problems like recognition of standard gestures and simple behaviors. Some progress has been made in building the statistical models of human behaviors using machine learning. However behavior recognition is complex, as the same behavior may have several different meanings depending upon the scene and task context in which it is performed. An alternative approach can be of providing *selective focus-of-attention* to the human supervisor by discriminating unusual or anomalous event from normal ones. Use of audio is encouraging in this respect because in many cases when the visual information required for a particular task is extremely subtle, audio stimulus is extremely salient for a particular anomaly detection task. Moreover, audio features can help in mining interesting patterns from the scene. Multimedia data mining has been applied for detection of events in sports video (like goal event in soccer [25] etc.) but it has not been systematically applied for surveillance videos. This requires development of a novel framework of low level feature extraction, advanced temporal analysis and multimodal data mining methods.

An obvious requirement for surveillance system is real time performance. If the system has to process signals from multiple sensor and modalities, then the required processing is multiplied. Moreover requirement for robustness and accuracy tend to make the algorithm design complex and highly computational. Generally commercial products employ embedded signal processing devices and high performance dedicated processors for faster processing but simultaneously increase the system cost heavily. Past research have focused on optimizing the low level image processing algorithms, reducing the feature space etc and recently researches on distributed surveillance system have tried to distribute the processing on the network. However there is lack of any systematic design and real time implementation of video processing algorithms using a network of multiple processors. Recent research and development in Grid technology can provide an alternative architecture for such implementation and research needs to be done to explore the feasibility of such innovative approach.

7 References

[1] DataMonitor. Global digital video surveillance markets: Finding future opportunities as analog makes way for digital. Market research report (July 2004). www.mindbranch.com/products/R313-6950.html.

[2] T. Bodsky, R. Cohen, E. Cohen-Solal, S. Gutta, D. Lyons, V. Philomin, and M. Trajkovic (2001). 'Visual surveillance in retail stores and in the home', *in: 'Advanced Video-based Surveillance Systems'*. Kluwer Academic Publishers, Boston, , Chapter 4, pp. 50–61

[3] J.M. Ferryman, S.J. Maybank, and A.D. Worrall (2000). 'Visual surveillance for moving vehicles', *Int. J. Comput. Vis.*, 37, (2), Kluwer Academic Publishers, Netherlands, pp. 187–19731

[4] R. Cucchiara, C. Grana, A. Patri, G. Tardini, and R. Vezzani (2004). Using computer vision techniques for dangerous situation detection in domotic applications. *Proc. IEE Workshop on Intelligent Distributed Surveillance Systems,* London, pp. 1–5.

[5] Frost & Sullivan report, (August 2005); www.frost.com

[6] W. Hu, T. Tan, L. Wang, and S. Maybank (August 2004). A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man and cybernetics*, 34(3):334–350,

[7] I. Haritaoglu, D. Harwood, and L. Davis. W4: Who, when, where, what: A real time system for detecting and tracking people. *In Third Face and Gesture Recognition Conference*, pp. 222-227.

[8] W. Niu, L. Jiao, D. Han, and Y. Wang (2003). Real-Time Multi-Person Tracking in Video Surveillance. *Proceedings of the Pacific Rim Multimedia Conference*, Singapore.

[9] A. Bobick and J. Davis (December 1996). Real time recognition of activity using temporal templates. IEEE workshop on Applications of Computer Vision, Sarasota, FL, 4 pages.

[10] J. Gao, A. G. Hauptmann and H. D. Wactlar. Combining Motion Segmentation with Tracking for Activity Analysis (2004). *The Sixth International Conference on Automatic Face and Gesture Recognition* (FGR'04), pp. 699-704, Seoul, Korea, May 17-19.

[11] E. Stringa,, and C.S. Regazzoni. Content-based retrieval and real-time detection from video sequences acquired by surveillance systems. *Int. Conf. on Image Processing, Chicago*, 1998, pp. 138–142

[12] J. Black, T. Ellis and P. Rosin. A novel method for video tracking performance evaluation (2003). *The Joint IEEE Int. Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, October, France, pp. 125–132.

[13] R. Cucchiara (2005). Multimedia surveillance systems. In *VSSN 05: Proceedings of the third ACM international workshop on Video surveillance & sensor networks, New York NY, USA*, pages 3-10.

[14] C. ´O Conaire, E. Cooke, N. O'Connor, N. Murphy, and A. F.Smeaton (2005). Fusion of infrared and visible spectrum video for indoor surveillance. *In International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS),* Montreux, Switzerland, April.

[15] C. stauffer. Automated Audio-Visual Activity Analysis. *CSAIL Technical Reports*, MIT. http://hdl.handle.net/1721.1/30568

[16] A. M. McIvor (2000). Background subtraction techniques. In *Image and Vision Computing, Hamilton, New Zealand*, Nov.

[17] J. J. Wang and S. Singh (2003). Video analysis of human dynamics – a survey. *Real-Time Imaging 9(5):* 321-346.

[18] R. Chellappa, C. L. Wilson and S. Sirohey (1995). Human and machine recognition of faces: a survey. *Proc. of the IEEE*, vol. 83, No. 5, pp705-740.

[19] V. I. Pavlovic, R. Sharma and T. S. Huang (1997). Visual interpretation of hand gestures for human computer interaction: a review. *IEEE Transactions on PAMI*, vol.19, no.7, pp.677-695, July.

[20] A. R. Dick and M. J. Brooks (2003). Issues in Automated Visual Surveillance. DICTA: 195-204.

[21] M. Valera and S.A. Velastin (2005). Intelligent distributed surveillance systems: a review vision. *In Image and Signal Processing, IEE Proceedings*, volume 152, pages 192 – 204, April.

[22] A. Amer and C. Regazzoni. *Editorial:* Introduction to the special issue on video object processing for surveillance applications. *Real-Time Imaging* Vol. 11, pp: 167–171, 2005.

[23] Nwagboso, C (1998). User focused surveillance systems integration for intelligent transport systems'*, in 'Advanced Video-based Surveillance Systems'*. Kluwer Academic Publishers, Boston, Chapter 1.1, pp. 8–12.

[24] R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y.Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L.Wixson (2000). A system for video surveillance and monitoring. *Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep*., CMU-RI-TR-00-12.

[25] M. Chen, S-C. Chen, M-L. Shyu, and K. Wickramaratna. Semantic Event Detection via Multimodal Data Mining. *IEEE Signal Processing Magazine*. pages 38-46, march 2006.

[26] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers (1999). Wallflower: Principles and practice of background maintenance. In *Proceedings of the Seventh IEEE International Conference on Computer IEEE Comput. Soc.*, volume 1, pages 255–261.

[27] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland (1997). Pfinder: real-time tracking of the human body. *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 780–785, July.

[28] C. Stauffer and W.E.L. Grimson (1999). Adaptive background mixture models for real-time tracking. In *Proceedings of CVPR99*, pages II:246–252.

[29] J. Barron, D. Fleet, and S. Beauchemin (1994) .Performance of optical flow techniques. *Int. J. Comput.Vis.*, vol. 12, no. 1, pp. 42–77.

[30] J.M. Ferryman, S.J. Maybank, and A.D. Worrall (2000). Visual surveillance for moving vehicles, Int. J. Comput. Vis., 37, (2), Kluwer Academic Publishers, Netherlands, pp. 187–197

[31] G. Welch and G. Bishop (2002). An Introduction to the Kalman Filter. *UNC-Chapel Hill, TR 95-041*.

[32] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp (2002). A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Trans. on Signal Process*., 50, (2), pp. 174–188

[33] A. Senior, A. Hampapur, Y.-L. Tian, L. Brown, S. Pankanti, and R. Bolle. Appearance models for occlusion handling. In *2nd IEEE Int. Workshop on PETS, Kauai, Hawaii, USA*, Dec 2001.

[34] B. Bhanu, I. Pavlidis, R. Hummel. *Guest Editorial:* Special issue on computer vision beyond the visible spectrum. Machine Vision and Applications (2000) 11: 265–266

[35] S.-S. Lin. Review: Extending visible band computer vision techniques to infrared band images. *Technical report, GRASP Laboratory, Computer and Information Science Department*, University of Pennsylvania, 2001.

[36] C. K. Eveland, D. A. Socolinsky, L. B. Wolff. Tracking human faces in infrared video. *Image and VisionComputing*, Vol. 21, pp. 579-590, 2003.

[37] M. Bertozzi, A. Broggi, P. Grisleri, T. Graf, M. Meinecke. Pedestrian Detection in Infrared Images. *Proc.IEEE Intelligent Vehicles Symposium 2003*, pp. 662-667, Columbus (USA), June 2003.

[38] F. Xu, X. Liu, and K. Fujimura. Pedestrian Detection and Tracking With Night Vision. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 6, No. 1, pp. 63-71, March 2005.

[39] J. Davis and V. Sharma. Robust detection of people in thermal imagery. In *Proc. Int. Conf. Pat. Rec.*, pages 713–716, 2004

[40] J. Han and B. Bhanu. Human activity recognition in thermal infrared imagery. *Computer Vision and Pattern Recognition*, 20-26 June, 2005.

[41] McDaniel, R., Scribner, D., Krebs, W., Warren, P., Ockman, N., McCarley, J. (1998). Image fusion for tactical applications. Proceedings of the SPIE - Infrared Technology and Applications XXIV, 3436, 685-695.

[42] J. Li. Spatial quality evaluation of fusion of different resolution images. *International Archives of Photogrammetry and Remote Sensing*, Vol.33, 2000.

[43] A. Toet. Heirarchical image fusion. *Machine Vision and Applications*, 3:1–11, 1990

[44] M. Pavel, J. Larimer, and A. Ahumada. Sensor fusion for synthetic vision. In *Conference on Computing in Aerospace, AIAA, 1991.*

[45] H. Li, B. Manjunath, and S. Mitra. Multisensor image fusion using the wavelet transform. In *Graphical Models and Image Processing*, volume 57, pages 234–245, 1995

[46] A. Utsumi, H. Mori, J. Ohya, and M. Yachida. Multipleview-based tracking of multiple humans. In *Proceedings of the 14th ICPR*, pages 597–601, 1998

[47] A. Nakazawa, H. Kato, and S.Inokuchi. Human tracking using distributed vision systems. In *Proceedings of the 14th ICPR*, pages 593–596, 1998

[48] H. Torresan, B. Turgeon, C. Ibarra-Castanedo, P. Hébert, X. Maldague,. Advanced Surveillance Systems: Combining Video and Thermal Imagery for Pedestrian Detection. In *Proc. of SPIE, Thermosense XXVI*, volume 5405 of *SPIE*, pages 506–515, April 2004.

[49] J. W. Davis and V. Sharma**.** Fusion-Based Background-Subtraction using Contour Saliency. *Computer Vision and Pattern Recognition*, 20-26 June, 2005

[50] C.O. Conaire, N. O.Connor, E. Cooke, A. Smeaton. Multispectral Object Segmentation and Retrieval in Surveillance Video. *To appear in International Conference on Image Processing,* 2006.

[51] P. Kumar, A. Mittal and P. Kumar. Fusion of Thermal Infrared and Visible Spectrum Video for Robust Surveillance. In *5th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)*, LNCS 4338, pp. 528 – 539, 2006.

[52] Ó. Conaire, C. O'Connor, N.E. Cooke and E. Smeaton. Comparison of fusion methods for thermo-visual surveillance tracking. *In International Conference on Information Fusion,2006.*

[53] D. L. Hall and J. Llinas. An introduction to multisensor fusion. *In Proceedings of the IEEE: Special Issues on Data Fusion,* pages 85(1):6–23, January 1997.

[54] R. R. Brooks and S.S. Iyengar. Multi-sensor fusion: fundamentals and applications with software. Upper Saddle River, N.J. : Prentice Hall PTR, 1998.

[55] S. Nazarko. Evaluation of data fusion methods using kalman filtering and TBM. Maters thesis. University of Jyväskylä. 2002.

[56] Ph. Smets. The Transferable Belief Model for Quantified Belief Representation. *In Handbook of defeasible reasoning and uncertainty management* systems. Gabbay D. M. and Smets Ph. Eds. Vol. 1, Kluwer, Doordrecht, 1998, pg. 267-301.

[57] G. Loy, L. Fletcher, N. Apostolo, and A. Zelinsky. An adaptive fusion architecture for target tracking. In *IEEE International Conference on Automatic Face and Gesture Recognition (FGR)*, 2002.

[58] A. Ayoun, and P. Smets. Data association in multi-target detection using the transferable belief model. *International Journal of Intelligent Systems*, 16:1167-1182. 2001

[59] G. Powell, D. Marshall, R. Milliken, K. Markham. A Data Fusion System for Object Recognition based on Transferable Belief Models and Kalman Filters. *In `Proceedings of 7th International Conference on Information Fusion.* Sweden, Pp 54-61. 2004

[60] P.J. Escamilla-Ambrosio, N. Mort. A Hybrid Kalman Filter - Fuzzy Logic Architecture for Multisensor Data Fusion. *Proceedings of the 2001 IEEE International Symposium on Intelligent Control* , pp. 364-369 , 2001

[61] Z. Elouedi, K. Mellouli, P. Smets. Assessing sensor reliability for multisensor data fusion within the transferable belief model. *IEEE Transactions on Systems, Man and Cybernetics, Part B,* page 782-787. 2004

[62] L. Snidaro, G.L. Foresti, R. Niu, and P.K. Varshney. Sensor fusion for video surveillance. *Proceedings of the seventh international conference on information fusion,* Vol. 2, Stockholm, Sweden, June 28th-July 1st, 2004, pp. 739-746.

[63] G. Rogova and V. Nimier. Reliability in information fusion: Literature. Survey. *In Proceedings of 7th International Conference on Information. Fusion,* Sweden, pp. 1158-1165. 2004

[64] P. Kumar**,** A. Mittal and P. Kumar. A Multimodal Audio, Visible and Infrared Surveillance System (MAVISS). *In Proceedings of the 3rd IEEE International Conference on Intelligent Sensing and Information Processing (ICISIP)*, pp. 151-157, 2005.

[65] J. Hershey, H. Attias, N. Jojic, and T. Krisjianson. Audio visual graphical models for speech processing. *In IEEE International Conference on Speech, Acoustics, and Signal Processing* (ICASSP04), Motreal, Canada, May 2004.

[66] A. V. Nefian, L. Liang, X. Pi, X. Liu, and K. Murphy. Dynamic bayesian networks for audio-visual speech recognition. *In EURASIP Journal on Applied Signal Processing* (JASP02), 2002.

[67] H. J. Nock, G. Iyengar, and C. Neti. Speaker localization using audio-visual synchrony: An

empirical study. *In International Conference on Image and Video Retrieval* (CIVR03), 2003. pages 488-499.

[68] D. Gatica-Perez, G. Lathoud, I. McCowan, J. Odobez, and D. Moore. Audio-visual speaker tracking with importance particle filter. *In IEEE International Conference on Image Processing* (ICIP03), 2003.

[69] N. Checka, K. W. Wilson, M. R. Siracusa, and T. Darrell. Multiple person and speaker activity tracking with a particle filter. *In International Conference on Acoustics Speech and Signal Processing* (ICASSP04), 2004.

[70] H. J. Nock, G. Iyengar, and C. Neti. Assessing face and speech consistency for monologue detection in video. *In ACM Multimedia*, 2002.

[71] P. K. Atrey, M. S. Kankanhalli and R. Jain. Information assimilation framework for event detection in multimedia surveillance systems. *Special Issue on "Multimedia Surveillance Systems" in Springer/ACM Multimedia Systems Journal*, September 2006

[72] D. Whitney, and J. Pisano, TASC, Inc., Reading, Massachusetts. *AutoAlert: Automated Acoustic Detection of Incidents.* December 26, 1995.

[73] M. Cowling, R. Sitte. Comparison of Techniques for Environmental Sound Recognition. *Pattern Recognition Letters,* Elsevier Science Inc., Vol. 24, Issues 15, pp. 2895-2907, Nov. 2003.

[74] Proximity Sensors. http://www.machinedesign.com

[75] Z. Zhu, W. Li and G. Wolberg. Integrating LDV Audio and IR Video for Remote Multimodal Surveillance. *IEEE Workshop on Object Tracking and Classification In and Beyond the Visible Spectrum*, in conjunction with IEEE Conference on Computer Vision and Pattern Recognition, San Diego, CA, June 2005

[76] I. Ulrich and J. Borenstein, "The guidecane, applying mobile robot technologies to assist the visually impaired," IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 31, no. 2, pp. 131–136, 2001.

[77] S. Shoval, I. Ulrich, and J. Borenstein, Computerized Obstacle Avoidance Systems for the Blind and Visually Impaired. Intelligent Systems and Technologies in Rehabilitation Engineering, CRC Press, 2000.

[78] Caimi, F.M. Bailey, B.C. Blatt, J.H. Undersea object detection and recognition: the use of spatially andtemporally varying coherent illumination. OCEANS '99 MTS/IEEE. Riding the Crest into the 21st Century ,pages 1474-1479 vol.3

[79] Heckman, P., and Hodgson, "Underwater Optical Range Gating", IEEE Journal of Quantum Electronics, QE-3, 11, Nov. I967

[80] A. F. Bobick, and J. W. Davis. The Recognition of Human Movement Using Temporal Templates. *PAMI*, Vol. 23, No. 3, 2001.

[81] K. Takahashi, S. Seki, H.Kojima, and R. Oka, "Recognition of dexterous manipulations from time varying images," in *Proc. IEEEWorkshop Motion of Non-Rigid and Articulated Objects*, Austin, TX, 1994, pp. 23–28.

[82] A. F. Bobick and A. D.Wilson. A state-based technique to the representation and recognition of gesture. *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 1325–1337, Dec. 1997.

[83] R. Duggad, U. B. Desai. A Tutorial on Hidden Markov Models. Technical Report No. SPANN-96.1. 1996. Indian Institute of Technology Bombay.

[84] L. Rabinier, A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. of IEEE*. 77 (2) (1989) 257-285.

[85] M. Leo, P. Spagnolo, T. D'Orazio, A. Distante. Human Activity Recognition in Archaeological Sites by Hidden Markov Models. *PCM (2)* , 1019-1026, 2004

[86] D. Mahajan, N. Kwatra, S. Jain, P. Kalra, S. Banerjee. A Framework for Activity Recognition and Detection of Unusual Activities. 15-21, *ICVGIP*, 2004, Kolkata, India.

[87] M. Yang and N. Ahuja. Extraction and classification of visual motion pattern recognition. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1998, pp. 892–897.

[88] U. Meier, R. Stiefelhagen, J. Yang, and A.Waibel. Toward unrestricted lip reading. *Int. J. Pattern Recognit. Artificial Intell.*, vol. 14, no. 5, pp. 571–585, Aug 2000

[89] O. Boiman and M. Irani. Detecting irregularities in images and in video. In *Proc. IEEE International Conference on Computer Vision*, pages 1985–1988, Beijing, China, Oct. 15-21 2005.

[90] D. Zhang, D. Gatica-Perez, S. Bengio, and I. McCowan. Semi-supervised adapted HMMs for unusual event detection. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, volume 1, July 2005.

[91] P. Kumar, A. Mittal and P. Kumar. A Multi-modal Data Fusion Framework Using Transferable Belief Model and Kalman filter for Robust Tracking in Dynamic Environment. Communicated to *Signal, Video and Image processing Journal,* Springer Verlag publication.

Extracting Named Entities and Relating Them over Time Based on Wikipedia

Abhijit Bhole
Indian Institute of Technology, Bombay, Mumbai - 400076, India
E-mail: abhijit.bhole@cse.iitb.ac.in

Blaž Fortuna, Marko Grobelnik and Dunja Mladenić
Jozef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
E-mail: {blaz.fortuna,marko.grobelnik, dunja.mladenic}@ijs.si

Keywords: text mining, document categorization, information extraction

Received: August 14, 2007

This paper presents an approach to mining information relating people, places, organizations and events extracted from Wikipedia and linking them on a time scale. The approach consists of two phases: (1) identifying relevant pages - categorizing the articles as containing people, places or organizations; (2) generating timeline - linking named entities and extracting events and their time frame. We illustrate the proposed approach on 1.7 million Wikipedia articles.

Povzetek: Predstavljene so metode rudarjenja informacij iz Wikipedie in urejanje v časovno zgradbo.

1 Introduction

Wikipedia is an abundant and valuable source of information manually constructed and mainly targeting human readers, and hence remains unfriendly towards automatic information extraction and mining. The text in Wikipedia is written using a special markup which is mainly aimed towards formatting the text and to a limited extent standardizing pages belonging to same categories. In this paper we propose an approach to extracting information from Wikipedia based on a standard method of first identifying pages that are relevant for the information extraction task and then extracting the desired information as illustrated in Figure 1. We use machine learning methods to identify pages belonging to the same category (in our case person, place or organization) as we have described in [12] and then proceed with text mining of articles to get links and time line information on named entities. The result of our approach is a collection of pages belonging to the predefined categories and a dynamic graph showing named entities (people, places and organizations) and relations between them. For instance, for each person we have important dates from his/her life and some events including places and organizations possibly associated with some dates. Our work is based on standard machine learning and text mining methods [7], in particular for document categorization we use linear support vector machine (SVM) [4], as it is currently considered the state-of-the-art algorithm in text categorization. We used binary linear SVM, the implementation from TextGarden [5]. The main novelty in this work is in extracting dynamic graph relating named entities based on Wikipedia. The closely related work is on semantically annotated snapshot of the English Wikipedia [1]. However that approach is based on natural language processing, while we are using machine learning and mining the extracted data in order to connect the extracted named entities. Another related work known as dbPedia [3] is on extracting structured information from Wikipedia mainly by using the existing structured information, such as Wikipedia *infobox templates* and by making this information available on the Semantic Web .

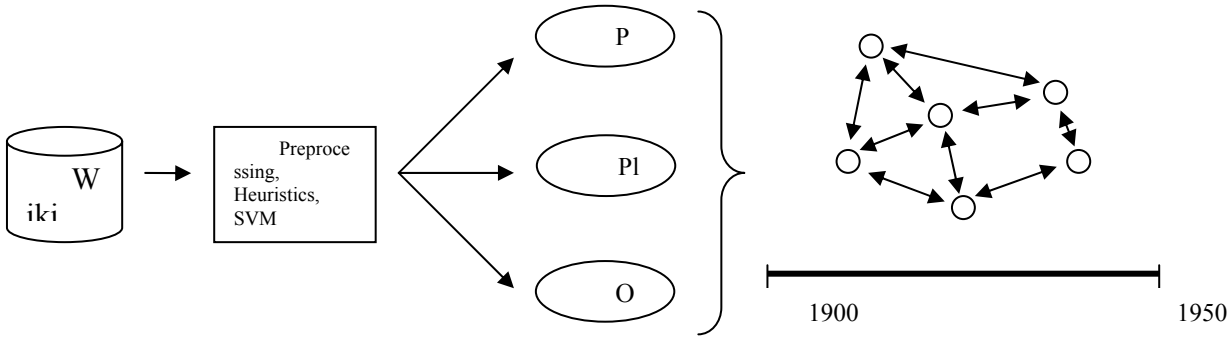


Figure 1. Illustration of the approach consisting of two phases: identifying pages containing named entities and generating timeline.

The task of building timelines from Wikipedia data was first pioneered by the project Wikistory [9], a project which creates a dynamic timeline of scientists based on the data obtained from dbPedia. Somewhat related is also work on Semantic MediaWiki which introduces some additional markup into the wiki-text which allows users to manually add "semantic annotations" to the Wiki [2].

The rest of the paper is structured as follows. Section 2 describes the proposed approach. Section 3 gives experimental results and illustrative example of the obtained dynamic graph. Section 4 concludes with discussion.

2 Approach description

Our goal is to generate a dynamic graph with named entities and relations between them, such as; a person was born at date X at city of Y, in year Z he/she lived in city W. In order to extract named entities we use Wikipedia and first categorize all its pages to find those containing the targeted named entities. We use two approaches to categorizing the pages: one based on manually constructed heuristic rules and the other based on labeling some of the examples and training an SVM classifier.

As Wikipedia pages contain articles with text and some markup information, there are two ways of approaching categorization of documents using heuristics: based on markup and based on text. We can analyze Wiki markup in the articles which can lead to clues for document categorization. However, we should be aware of the fact that the Wiki markup is not semantic annotation and is not uniformly and strictly enforced across the whole Wikipedia. The other approach is to extract plain text from Wiki and then search for particular phrases. We use regular expression based search and replace method throughout for parsing the Wiki markup, obtaining plain text as well as searching for heuristics.

Categorization of Wikipedia pages can be realized either by using index pages on Wikipedia or by classifying each individual articles. Our preliminary experiments using index pages show that it is not a reliable and consistent method over the long term since the index pages are not automatically generated but are maintained by the users themselves. We also tried with filtering some page titles based on tagged word corpuses, but the results were of limited value. Our current strategy involves analyzing individual articles by first using heuristic-based categorization of articles and then, once we have a sufficient number of articles, applying machine learning techniques. A brief description of these heuristics is given along with the experimental results in Section 3.

Once we have categorized Wikipedia pages as describing people, place or organization getting the corresponding named entities is trivial from the Web page URL (e.g., Wikipedia page describing Abraham Lincoln is on URL http://en.wikipedia.org/wiki/Abraham_lincoln). Once we have a collection of articles describing named entities we apply mining for relations between the extracted named entities. We define relation between two entities as follows. An entity is related to other if it has a certain probability of reaching that entity using the Wikipedia hyperlink structure. This usually reflects real life relationships, but inevitably with some difference. The approach we use for finding relations is to build a matrix based on out-links and in-links from a Web page and then searching for two named entities occurring together in the text of other articles.

There are two possible ways of realizing events: one is to identifying articles belonging to events and other is to apply text mining on article texts searching for sentences having mention of dates in them. The task here involves first sentence boundary disambiguation, then extracting dates from these sentences and also linking them to the corresponding named entities.

3 Experiments

3.1 Overview

The entire Wikipedia consists of 5 million entries including redirections, list pages and namespace pages (“Category:”, “Template:” etc), out of which about 1.7 million can be considered as articles describing some unique concept. These 1.7 million articles can be obtained by simple heuristic filtering, which forms our main corpus. Out of that corpus, we selected a random sample of 1000 articles and manually labeled them with three categories (describing people, place, and organization) and left the rest unlabeled. In that sample of articles we found 260 people, 184 places and 118 organizations. We used that for evaluating our approach and estimating the total number of named entities present in the corpus. Although the size of the manually labeled test sample is rather small, it is our hypothesis that it represents the diversity of the source.

Our preliminary experiments identifying named entities through index pages were made by crawling pages with titles similar to “list of people” resulting in about 130K hits which is estimated to result in about 30% recall and about 85% precision. An effort to filter out titles using word corpus failed to have any significant impact on accuracy due to the fact that many persons have fairly common words in their names while some non-person articles may have uncommon words. Similar was true for the other two categories (places, organizations).

3.2 Heuristic based identification

We use parts of Wiki markup known as “Infobox” for identifying some people, places and organizations. We then search for articles with longitude and latitude co-ordinates and annotated them as place while we use the first paragraph from the extracted plain text from the article to search for keywords and dates for identifying people. In this way we labeled 285.000 articles as people, 150.000 as places and 26.000 as organizations. We have evaluated the proposed heuristic approach on our

manually labeled sample of 1000 articles. The results are summarized in Table 1.

	Precision	Recall
People	100%	62.4%
Places	95.7%	48.9%
Organizations	100%	10.16%

Table 1: Evaluation of heuristics based classification of Wikipedia articles.

More detailed analysis of the obtained results shows several common sources of error. Many people left out by our heuristic do not have indicative dates in sufficient proximity of the beginning of article. Also the heuristic takes into account most common formats of date, but is conservative enough not to include irrelevant text such as ISBN numbers to maintain its precision. After all, the goal is to apply machine learning with sufficient labeled data rather than to rely entirely on heuristic. When identifying places, several articles on military vehicles or asteroids may have co-ordinates in them but do not qualify for this category which leads to the error. An organization is a loosely defined term, can be any entity which links people (company, university, school or even a rock band). Hence they lack any such indicative text and remain difficult to be found simply by this technique. Some examples of successfully identified named entities are given in Figure 2.

"Aristotle (Greek: Ἀριστοτέλης *Aristotélēs*) **(384 BC – 322 BC)** was a Greek philosopher, a student…"

Example 1: The article describing Aristotle which belongs to category person having birth dates at the beginning of article.

{{Infobox_**Philosopher**
| me = {{polytonic|Ἀριστοτέλης}} "Aristotélēs"
| image_name = Aristoteles Louvre.jpg
| color = #B0C4DE
| region = Western philosophy
| era = [[Ancient philosophy]]
| name = [[Aristotle]]
| birth = [[384 BC]]

Example 2: The article describing Aristotle has a specific Wiki markup indicating him to be a philosopher and thus classifiable as a person.

The coordinates of the nominal centre of London (traditionally considered to be the original [[Eleanor Cross]] at [[Charing Cross]], near the junction of [[Trafalgar Square]] and [[Whitehall]]) are approximately **{{coordms|51|30|29|N|00|07|29|W|type:city(7,000,000)_region:GB}}.**

Example 3: A co-ordinate header in the article describing London.

Figure 2. Examples of Wikipedia text of some identified named entities for category person (Examples 1 and 2) and place (Example 3).

3.3 Machine learning based identification

In order to apply machine learning, text of the documents to be classified was preprocessed by removing stop-words applying stemming and representing each document using the standard bag-of-words approach containing individual words. Representation of each document was further enriched by frequent phrases, which were considered to be those consisting of up to two consecutive words [6] and occurring at least fifty times in the data collection. The binary classification model was automatically constructed using Support vector machines for each of the three classes, taking the training documents of the class as positive and the training documents of other classes as negative. The classification model was trained on one part of the data collection, leaving the other part to be classified using the standard statistical method called cross validation. In other words, the data collection was randomly divided into several disjoint parts (in our case three) of approximately equal size. Then three classification models were generated, each taking one of the three parts as testing and the remaining two parts as training documents. We report average performance (precision, recall) over the three models. We first ran 3-fold cross validation on manually labeled dataset, using only the first paragraph from the article text (Table 2a) and using the plain text from the entire article (Table 2b).

	Precision	Recall	F1	BEP
People	92.81%	49.20%	64.22%	60.40%
Places	73.14%	54.46%	62.03%	64.88%
Org.*	25.96%	49.17%	33.91%	28.05%

Table 2a: Results from cross validation on sample set using only first paragraphs of article text.

	Precision	Recall	F1	BEP
People	85.18%	63.38%	72.67%	75.57%
Places	85.31%	60.29%	70.46%	77.02%
Org.*	47.63%	37.73%	41.59%	43.11%

Table 2b: Results from cross validation on sample set using plain text from entire article.

*Since the sample set was too unbalanced in this category, the cross validation was run with bias misclassification cost (SVM parameter j=5).

The text extracted from the first paragraph of the articles was used in all subsequent classification experiments presented in this paper since it is significantly computationally less expensive. However, the results obtained using the entire text of the articles (Table 2b) are better than the results of using the first paragraph only (Table 2a). We also tried to use the text including the Wiki markup and it produced results that

are a few percent better than those in Table 2b, but this needs further analysis.

The diversity of the sample set can be explained to be a cause of the low recall in case of people and places, where the SVM may have misclassified some article alien to it. However, the precision of people and places was encouraging considered the small size of the sample set as the SVM succeeded in picking up most of the entities which were sufficiently represented in this training data. The poor results with organizations can be explained either by insufficiency of the sample set to capture the diversity of the category or by inability of the bag-of-words text representation to capture all the features necessary for identifying organizations.

We also ran 3-fold cross validation using only the first paragraphs on the whole corpus, which was labeled solely by our heuristics. In the first experiment we use a binary SVM to train the classification model (Table 3a) and in the second experiment we use one class SVM (Table 3b).

	Precision	Recall	F1	BEP
People	83.59%	79.72%	81.61%	82.30%
Places	83.58%	72.19%	77.47%	78.58%
Org.	38.85%	58.28%	46.62%	45.66%

Table 3a: Results of cross validation using binary SVM on the whole corpus

The experiments show that binary SVM is much more successful than one class SVM in partitioning our data. Also, we can assert that our hypothesis that the features present in the text are good enough to enable automatic classification of named entities is correct to a large extent.

	Precision	Recall	F1	BEP
People	46.66%	89.50%	61.34%	63.74%
Places	38.01%	89.39%	53.34%	51.88%
Org.	9.41%	86.32%	16.97%	23.72%

Table 3b: Results of cross validation using one class SVM on the whole corpus.

When running binary SVM on the entire text of articles, we didn't observe considerable improvements over the results given in Table 3a. Namely, we got F1 of 79.37%, 78.79% and 37.09% for people, places and organizations respectively and BEP of 80.14%, 80.05% and 52.01%.

3.4 Analyzing Relations

We analyzed relationships between the total of 285,549 persons for three different kinds of clues relating them: in-links, out-links occurring in the article pages and persons mentioned in the same paragraph of a text of another article or a list. Each of the three clues was given a weight based on preliminary experiments. Then an adjacency matrix ("to-from") of a graph with persons as nodes was constructed and the end result was obtained by sorting the edges by their weight. Further extension is possible by recursively multiplying the matrix in order to consider paths having length more than one [8]. Example result providing relations for Aristotle is given in Figure 3.

In our experiment, about 2.95 in-links / out-links per person were taken into consideration. Apart from that references to pairs of persons occurring together across total 5,735,346 pages were used to assign weights to relations. Figure 4 gives percentage of the number of relations over different weights larger then 0.1. We can see that 46% of relations have weight between 0.1 and 0.2 (578,083 out of 1,258,807).

In addition to associations based on explicit links and co-occurrences, further association is also possible by finding overlaps between time lines of two people. Preliminary analysis on the sample of data has shown that the proposed approach is promising; however a quantitative analysis of these results is part of the future work.

Aristotle is related to:

Plato	(25.217420)
Thomas Aquinas	(4.700384)
Socrates	(4.536786)
Cicero	(3.608422)
Alexander the Great	(3.017379)
Plutarch	(3.011533)
Averroes	(3.000203)
Demosthenes	(2.028392)
Ptolemy	(1.938013)
Aristophanes	(1.848224)
Avicenna	(1.823166)
Galileo Galilei	(1.714287)
Hippocrates	(1.688921)
Euclid	(1.670485)
Homer	(1.659085)

Figure 3: The first 15 persons related to Aristotle are shown along with their corresponding importance weights as suggested by our algorithm.

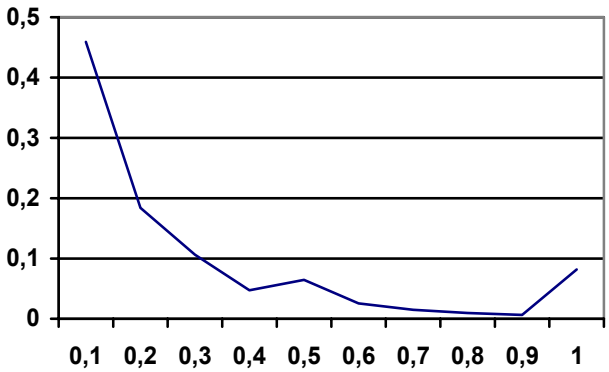


Figure 4: Distribution of relations across different weights.

3.5 Events extraction

To extract events, we used heuristic-based sentence boundary disambiguation after extracting plain text from

Wiki markup and then picked up sentences containing dates. The extracted sentences are regarded as events and are linked to the article in which they were found (see Figure 5 for an example). Other named entities are searched for in the extracted sentences. However the task of incorporating these entities into relationships was already accomplished previously.

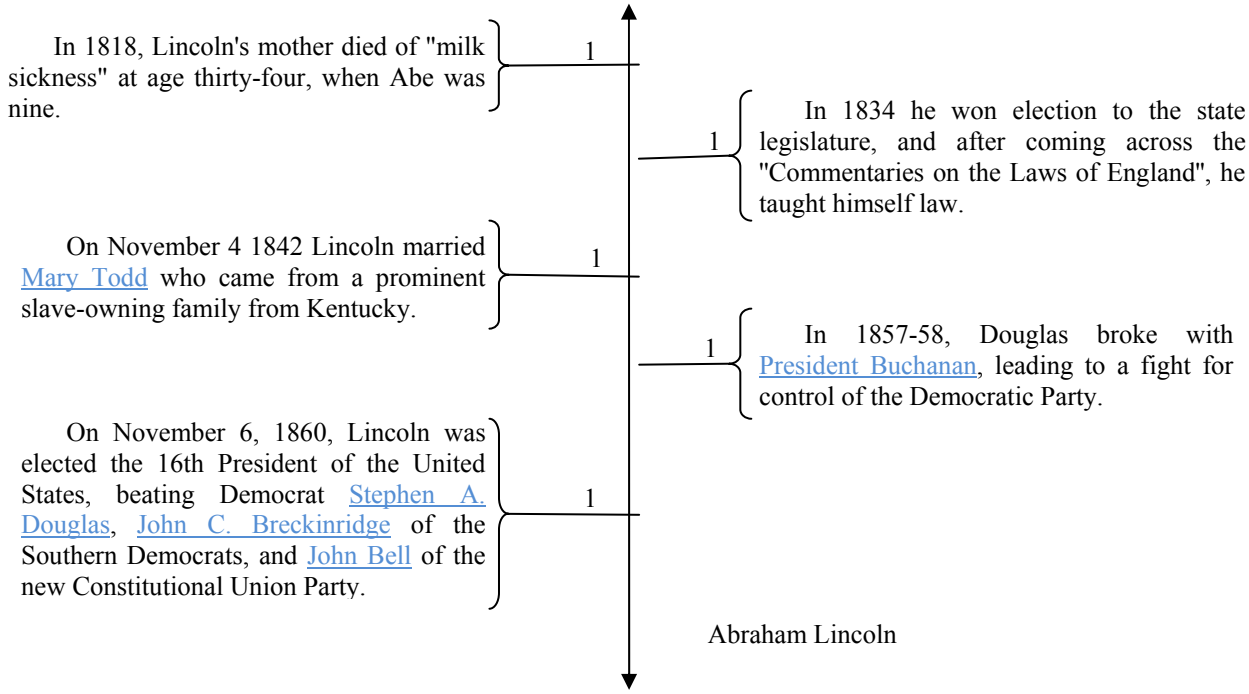


Figure 6: Illustration of timeline generated for Abraham Lincoln showing connections to other identified persons.

Overall we have extracted 2,148,602 events making in average 7.5 per person. In addition we extracted 270,969 birth dates and 113,374 death dates.

- 0/0/-323 - Upon Alexander's death in 323 BC, anti-Macedonian feelings in Athens once again flared.
- 0/0/-86 - When Lucius Cornelius Sulla occupied Athens in 86 BC, he carried off the library of Appellicon to Rome, where they were first published in 60 BC by the grammarian Tyrranion of Amisus and then by philosopher Andronicus of Rhodes.

Figure 5: Example events extracted from the article describing Aristotle.

For the people that we did not manage to obtain birth and death date we used heuristic estimating the time period of the person based on the dates occurring in the events. This heuristics enabled association of named entities based on time lines.

Once we have extracted events and linked them with named entities we generate timeline for a selected named entity that is linking it with other named entities and showing the time frame of their interactions. Illustration of an example generated timelines given in Figure 6.

4 Conclusions

In this paper we outlined how heuristic based approaches can be used for extracting high quality annotations of Wikipedia articles and that automatic text categorization is a viable way of generalizing the heuristics. We have proposed an approach to that consist of two phases: (1) identifying relevant pages containing people, places or organizations and (2) generating timeline linking named entities via the extracting events and their time frame. As a part of future work we are planning to further extend the number of extraction classes, detail level of extracted data from the articles (e.g. type of place, profession of a person, etc.) and investigate integration of the extracted facts into a knowledge base, such as Cyc [10, 11].

Acknowledgement

This work was partially supported by the Slovenian Research Agency and the IST Programme of the EC under NeOn (IST-4-027595-IP) and PASCAL (IST-2002-506778).

References

[1] H. Zaragoza and J. Atserias and M. Ciaramita and G. Attardi. Semantically Annotated Snapshot of the English Wikipedia v.0 (SW0), http://www.yr-bcn.es/semanticWikipedia, 2007

[2] Semantic MediaWiki, http://ontoworld.org/wiki/-Semantic_MediaWiki, 2007

[3] DBpedia.org, http://dbpedia.org/docs, 2007

[4] Joachims, T. Making large-scale svm learning practical. In B. Scholkopf, C. Burges, and A.

Smola, editors, Advances in Kernel Methods, Support Vector Learning, MIT-Press, 1999.
[5] Grobelnik, M., Mladenic, D. 2007. Text Mining Recipes. Berlin, Heidelberg, New York: Springer. Accompanying software available at: http://www.textmining.net.
[6] Mladenic, D., Grobelnik, M. 2003. Feature selection on hierarchy of web documents. Journal of Decision Support Systems, 35. 45-87.
[7] Sebastiani, F. 2002. Machine learning in automated text categorization. ACM Computing Surveys, 34(1). 1–47.
[8] Olston, C., Chi, H. E. (2003). ScentTrails: Integrating browsing and searching on the Web. In ACM Transactions on Computer-Human Interaction (TOCHI), Volume 10, Issue 3, Pages: 177–197, ACM Press, New York, NY, USA.
[9] Wikistory, http://www.urbigene.com/wikistory/
[10] Douglas B. Lenat, (1995). "Cyc: A Large-Scale Investment in Knowledge Infrastructure." Communications of the ACM 38, no. 11, November 1995.
[11] Purvesh Shah, D. Schneider, C. Matuszek, R.C. Kahlert, B. Aldag, D. Baxter, J. Cabral, M. Witbrock, J. Curtis, (2006). Automated Population of Cyc: Extracting Information about Named-entities from the Web. In Proceedings of the Nineteenth International FLAIRS Conference, pp. 153-158, Melbourne Beach, FL, May 2006.
[12] Bhole, A., Fortuna, B., Grobelnik, M., Mladenić, D. Mining wikipedia and relating named entities over time. In: Bohanec, M., Gams, M., Rajkovič, V., Urbančič, T., Bernik, M., Mladenić, D., Grobelnik, M., Heričko, M., Kordeš, U., Markič, O. Proceedings of the 10th International Multiconference on Information Societz IS 2007, 8.-12. october 2007, volume A. Ljubljana: "Jožef Stefan" Institute, 2007, pp. 177-180.

Intelligent High-Security Access Control

Matjaž Gams and Tea Tušar
Department of Intelligent Systems
Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
E-mail: matjaz.gams@ijs.si, tea.tusar@ijs.si

Keywords: access control, machine learning, intelligent data analysis, outlier detection

Received: July 12, 2007

Access control is an important security issue in particular because of terrorist threats. Access points are increasingly becoming equipped with advanced input sensors often based on biometrics, and with advanced intelligent methods that learn from experience. We have designed a flexible modular system based on integration of arbitrary access sensors and an arbitrary number of stand-alone modules. The system was tested with four sensors (a door sensor, an identity card reader, a fingerprint reader and a camera) and four independent modules (expert-defined rules, micro learning, macro learning and visual learning). Preliminary tests of the designed prototype are encouraging.

Povzetek: Članek opisuje vgradnjo inteligentnih metod v sistem za nadzor vstopa.

1 Introduction

Attacks on civil and institutional objects are becoming a potential threat in several parts of the world. The target might be a bank or a company, and the attack might be motivated by money or ideological reasons, but the essential pattern is the same. Due to the increase of these types of attacks it is important that advanced scientific and technological solutions are applied to real-life applications.

One of the important security tasks is to assure efficient access control, e.g. to differentiate between "proper" access of "fit" employees and all other attempts of access. An example of such an access point is presented in Figure 1.

Figure 1. A case access point: the door opens when two sensors on the right side confirm the identity of the acceding person. The upper sensor is an identity card reader and the lower a fingerprint reader.

These days, protective security uses a multi-layered approach, known as *defence in depth*. Defence in depth means combining several measures to make unauthorized access difficult for an external intruder or an employee, which is either without the needed permissions or not in "the right state of mind". In the defence in depth concept, various security measures complement and support one another, including physical space, security procedures, personnel and technology. Sub-concepts include also security awareness based on the co-operation of staff that fully know their responsibilities, and include human guards and security teams at various levels.

Quite often, the security system already includes one or more biometric sensors (Kolbe and Gams 2006; Ashbourn 2003; Jain et al. 1999). According to M. Kirkpatrick, assistant director of the FBI's criminal justice services division, "the only way to trace a terrorist is through biometrics". Indeed, biometrics is hard to overcome by itself (Wayman 2004; Toledano et al. 2006; Lumini and Nanni 2006). But as can be found on the Internet by simply browsing YouTube (for example search: "MythBusters beat fingerprint security system"), each method can be fooled quite easily once the security mechanism is figured out. Therefore, in this paper we are interested in the introduction of intelligent methods into an existing security access system, thus adding another layer of security.

Intelligence is generally accepted as one of the most important factors when fighting crime and terrorism, but the term is usually limited only to human intelligence. In this paper, we are concerned with intelligent computer methods (Hopgood 2000; Albus and Meystel 2001; Turban et al. 2004) that learn from past experience and react on the basis of the so obtained knowledge (Mitchell 1997; Witten and Frank 2005).

Intelligent methods can be used to monitor the behaviour of people at an access point on two levels. On

micro level, we take advantage of the fact that at an access point, a person always acts in a similar manner that rarely changes over time and is usually different for different people. The way a person accesses the identification and verification sensors depends on his/hers habits and motoric abilities. For example, person #1 always carries his identity card in his wallet and puts the whole wallet near the identity card reader, while person #2 keeps her identity card in the purse and always needs some time before pulling out her identity card and putting it back into place after identification. Similarly, some people throw open the door, while others open the door just enough to slip through. In short, because of people's usual behaviour on the micro level we can learn their pattern and use this information to further verify the accessing person.

Behaviour on the *macro level* refers to the daily routine of the users of the access control system. While a person can have the required permissions to all access points in the system, it usually accesses only some points or the access depends on the current week-day or hour. Similarly, smokers usually exit the building more often than non-smokers etc. The system can also detect dependencies between users, such as person #1 and person #2 always enter or exit at the same access point in a short period of time.

In this way the intelligent module acts as a security guard that is familiar with each employee and notices not only clear violations of attempted access by unauthorised people, but also abnormal states of the employees, e.g. in cases of drug abuse or mental imbalance. Additionally, the learnt patterns can be later analyzed by the security personnel.

Machine learning methods are among the most popular artificial intelligence techniques and have been used in various applications. Several open-source machine learning toolkits are available on the internet, in particular Orange (Demšar et al. 2004) and Weka (Witten and Frank 2005). It is no surprise that machine learning methods are also used for security tasks.

The rest of the paper is structured as follows: Section 2 reviews the related work. The structure of the proposed system is presented in Section 3, while Section 4 is dedicated to the sensor level of the system. In Section 5 we show how machine learning methods can be used for solving this problem. Details on the used modules are given in Section 6. Finally, Section 7 is dedicated to the experimental evaluation of the presented system, while Section 8 concludes the paper by summarizing the work done.

2 Related work

Existing access control systems usually consist of a combination of identification and verification modules that range from traditional knowledge-based (e.g. passwords or PIN codes) and token-based (e.g. ID cards and smart cards) security modules to more advanced sensors that are based on biometric information from handwriting, fingerprint, face, voice, retina, iris, hand geometry and even vein patterns. See (Kolbe and Gams 2006; Li et al. 2006) for a review of existing biometric sensors applied to access control. Additionally, the most advanced control systems rely on video cameras to monitor the behaviour of users. The so-called *intelligent video systems* (Wilson 2005) enable differentiation between normal and abnormal behaviour of users and are often capable of covering a wide perimeter.

However, the mentioned systems do not implement a higher level of intelligence that would serve as an additional security module. While in (Lamborn and Williams 2006) machine learning methods are used to combine the results of heterogeneous sensors into a global solution, to our best knowledge there is no working access control system that applies machine learning methods to detect abnormalities in user behaviour on micro and macro level. Instead, machine learning is often used on similar applications, such as keystroke dynamic user authentication (Revett et al. 2007), computer network intrusion (Eskin et al. 2002; Lane and Brodley 1999) or intrusion detection in web applications (García Adeva and Pikatza Atxa 2007). The applied algorithms range from probabilistic neural networks and support vector machines to *k*-nearest neighbour and cluster-based algorithms.

3 Structure of the system

To design and test our intelligent methods for access control, we have set up an experimental environment, as presented in Figure 2. It consists of a single access point with a door equipped with an open/close sensor, an identity card reader, a fingerprint reader and a camera. After the two readers have successfully identified and verified the user, the door unlocks. The user opens the door, passes through the door and the door automatically closes. The whole entry process is recorded by the camera, while the input signals of the other three sensors are connected using a controller and inputted into a database. For each successful access, the following four times are registered:

- time of acceptance of the identity card,
- time of acceptance of the fingerprint,
- time of door opening,
- time of door closing.

For the proposed system, the number of registered times is adaptable to a particular configuration, but should not be too small, e.g. less than three, or too big, e.g. hundreds.

After the user passes the access point, the data collected during the access is additionally processed by the four modules of our intelligent system:

- expert-defined rules (cover the unwanted behaviour on the macro level),
- micro learning (learns the patterns of access on the micro level),
- macro learning (learns the patterns of access on both the macro and micro level – it consists of three sub-modules that are presented in more detail in Section 6),
- visual learning.

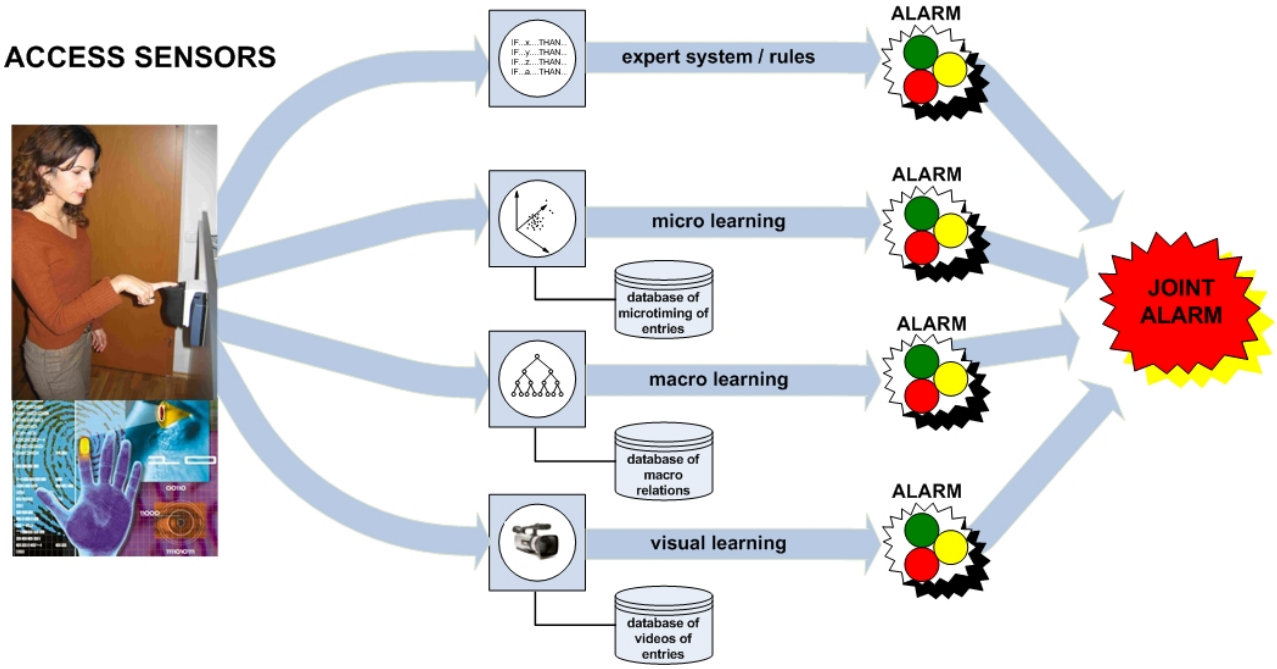


Figure 2. The structure of the intelligent access control system prototype.

Each module performs classification on its own and the final classification is calculated from the basic modules. Each module can classify an access into one of the following three categories:

- OK (according to the module, the access is regular),
- warning (it is unclear whether the access is regular or not),
- alarm (according to the module, the access is irregular).

An example of the system's output is presented in Figure 3. All the texts are in Slovene. Red color represents alarm. The macro module classification consists of the classifications of its three sub-modules.

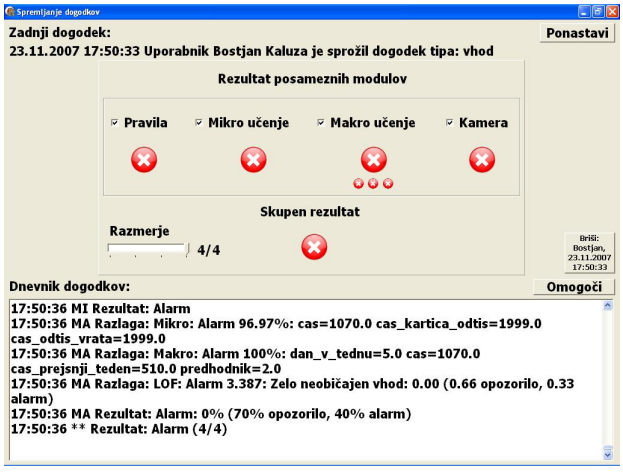


Figure 3. The graphical user interface of our intelligent access control system for an irregular access. The four big circles represent classifications of the four stand-alone modules. The single circle bellow represents the joint and final classification, while the text at the bottom shows the explanations for each module.

All modules perform in parallel and each transmits its decision and explanation when it finishes. Therefore, the classifications in the graphical user interface appear asynchronously. The log of module events consists of the exact time of module decision, followed by its classification, certainty factor and additional explanation of the decision (where possible).

4 Access sensors

The four sensors used in our experiments (door sensor, identity card reader, fingerprint reader and camera) are standard commercially available sensors that do not need additional explanation. While the two used readers allowed only the time of acceptance to be recorded, different identification/verification devices allow the recording also of other features, e.g. time of start of the identification/verification process, confidence rate, time between pressing buttons of a console etc. Our approach is in general independent on the number of such features, but all the experiments were performed using the mentioned four sensors and four recorded times.

4.1 The DOX controller

The input signals from the door sensor, identity card reader and fingerprint reader were collected by DOX – a multi-channel access controller, developed by the company Špica International (see Figure 4). DOX can be connected to various peripheral devices, such as card readers, touch screen consoles, door locks, biometric readers and other activators and sensors.

A single DOX controller is bound to one access point, however, multiple DOX controllers can be dynamically combined thus enabling the use of up to hundreds of input sensors on multiple locations in a

single access control network. For our purposes one DOX controller sufficed.

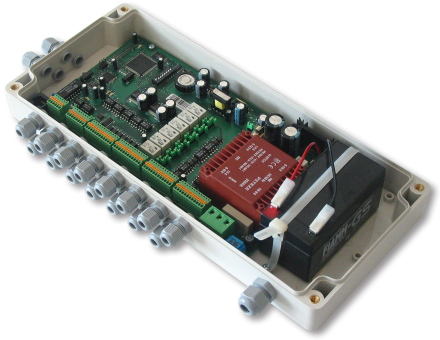


Figure 4. The DOX access controller.

4.2 Time&Space database

The Time&Space software is a commercial product by Špica International for access control and timekeeping. For the purposes of our research, the database of Time&Space needed minor modifications. In particular, a more detailed record of times from the input sensors was implemented. Essentially, all the other features, like scalability, modularity and security, remained the same thus enabling a reliable and flexible database.

5 Machine learning methods

Among the four independent modules of the intelligent system, only one (the expert-defined rules) does not use learning. The remaining three modules basically apply the general schema of machine learning, presented in Figure 5 (Kotsiantis 2007).

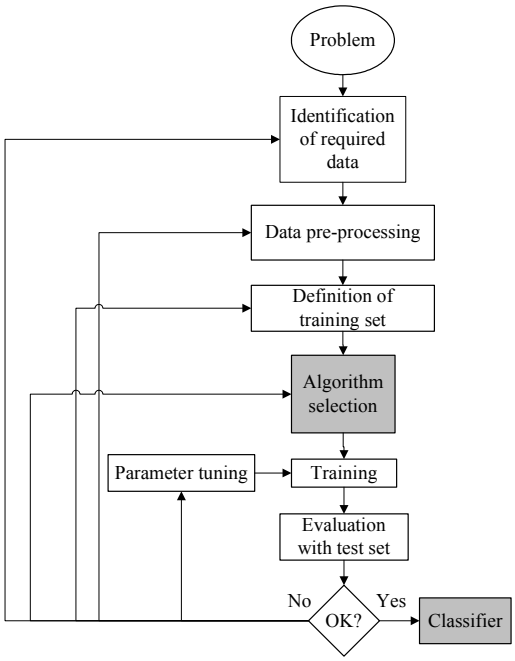


Figure 5. The process of applying supervised machine learning to a real-world problem (Kotsiantis 2007).

Inductive machine learning is the process of learning from instances (examples in a training set). As a result, a classifier is created that can be used to classify new instances. During the process, several steps are needed to achieve the desired functionality. For the security access control classification, the following properties of the classifier are the most important:

- the classifier should be as accurate as possible,
- the structure of the constructed knowledge should be comprehensible to the designers and system engineers as well as maintenance personnel,
- the classification result should be accompanied by an explanation that is comprehensible to all security personnel.

The first step in applying machine learning methods to a real-world problem is collecting the input data and features (Zhang et al. 2003). In our access control system, we had two major sources of input data and features – on the micro and macro level.

On the micro level, we record the video of the access together with times of identification, verification and door opening and closing. On the macro level, we note different macro features, such as time in the day, day of the week, persons that preceded the current person etc. The applied method of gathering features is a "brute-force" method, meaning that all the available features were gathered in the first stage.

The next step was gathering of data. In our experimental setting, only positive examples (regular accesses) were given as input data. Each access was recorded and stored into the database. It was then up to the individual modules to pre-processes the data for their specific needs.

Hodge and Austin (2004) have introduced a survey of contemporary techniques for outlier detection in the context of eliminating "bad" instances. However, in our case it was easy to visually eliminate the bad entries due to unexpected human errors during access. Namely, for high-security access, each person has to enter in a typical manner with as little variation as possible. If any unusual variation appears, the system is supposed to notice the deviation. A reasonable amount of robustness is necessary, but nothing more. In summary, it was not difficult to eliminate all but normal entries of all the tested persons. In a realistic application, it should not be difficult to eliminate those entries that stand out as outliers within the group of entries of one person.

The next step is feature selection – the process of identifying and removing as many irrelevant and redundant features as possible (Yu and Liu 2004). This reduces the dimensionality of the data and for large datasets enables machine learning algorithms to operate faster and more effectively. In our case there was no problem with too many features, but it is well known that typically some features depend on one another thus often unduly influencing the accuracy of supervised machine learning classification models. This problem is sometimes addressed by constructing new features from the basic feature set (Markovitch and Rosenstein 2002), which was also one of our intentions. However, the

relatively small changes we introduced were basically all hand-crafted.

On the basis of the created datasets, the choice of specific learning algorithms was performed. The algorithms were chosen from the machine learning toolkits Weka and Orange, both freely available from the internet. In addition, we have designed algorithms on our own. The first choice of algorithms was based on the task itself since it is well known that specific algorithms are suitable for specific types of problems. The second choice was based on algorithms' explanation capabilities. Two types of algorithms were finally chosen, one for constructing decision trees and one for outlier detection, as described in the next section.

6 Stand-alone modules

Our intelligent access control system consists of four stand-alone modules: expert-defined rules, micro learning, macro learning (combining the three sub-modules) and visual learning. They are presented in more detail in this section.

6.1 Expert-defined rules

Even existing commercial access systems use several rules to control basic behaviour on the macro level. Such rules describe, for example, the maximal time a door can be open or if a person can enter a building on a Saturday. In addition to taking into account these "common" rules we gathered new rules by consulting a security expert. Through knowledge acquisition we have obtained sufficient information to design altogether nine rule templates. Each rule template must be filled with exact data using a rule editor. As a result, several or several tens of concrete rules can be created. Some rules are quite simple while others demand specialized variables and routines to be performed properly.

An example of a rule template is: *The module triggers MESSAGE if user(s) SET_USERS do not exit the building in time TIME_LIMIT. MESSAGE* can be either a warning or an alarm, while *TIME_LIMIT* is an integer value measured in seconds. *SET_USERS* is an example of a special routine, which enables to choose a particular user, a particular group of users or all users of the system. This rule obviously demands a special hand-coded routine that checks all the persons that entered the buildings and did not exit and compares the time spent in the building with *TIME_LIMIT*.

The rules in the current implementation are not being chained as in a typical expert or a rule-based system. Rather, the inference engine checks all the rules on the list one by one and triggers messages in those rules that match the preconditions. Rules are transformed into SQL queries and are executed on the Time&Space database. If any rule triggers a warning (or an alarm), the classification of the whole module is warning (or alarm).

6.2 Micro learning

The micro learning module uses only the information gathered on the micro level. From the four input times recorded during each access (see Section 3), it calculates three time intervals, measured in microseconds:

- time between acceptance of the identity card and acceptance of the fingerprint,
- time between acceptance of the fingerprint and door opening,
- time between opening and closing of the door.

Each access can be thus represented as a point in a three-dimensional space defined by these three times (see Figure 6). All regular accesses of a person compose the training set for this person. When a new access by this person (or someone who impersonates this person) is made, the new three-dimensional point is compared to the existing "regular" points. If it does not match the regular accesses, the module classifies it as either a warning or an alarm (depending on how much it differs from the training set).

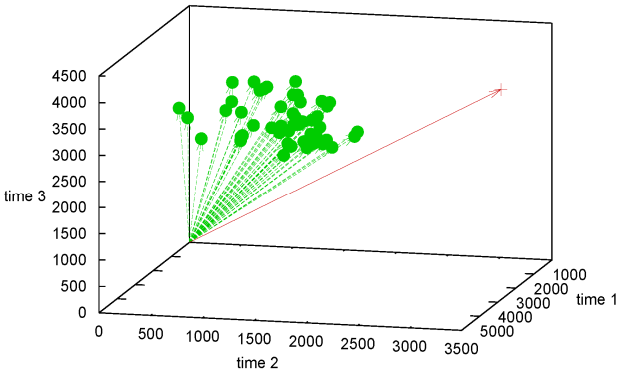


Figure 6: The three-dimensional representation of access times used in micro learning.

Outlier detection algorithms can be used for the purpose of micro learning. In (Tušar and Gams 2006) we reviewed existing outlier detection algorithms and chose the so-called Local Outlier Factor (LOF) for our task (Breunig et al. 2000; Breunig 2001). This algorithm was preferred to the others because it calculates a real number representing the "degree of outlierness" for every point (the greater the value of LOF, the more outlying the point). This gives us the possibility to define warnings and alarms by simply setting the bounds to the value returned by LOF. For example, taking the values 1.1 and 1.3 as bounds, all points that have the LOF value lower than 1.1 are classified as OK, the points that have the LOF value between 1.1 and 1.3 are classified as warning, while all the other points (having the LOF value greater than 1.3) are classified as alarm. While this gives us great flexibility, it also requires the proper setting of these bounds for the algorithm to work properly.

6.3 Macro learning

The macro learning module is composed by three sub-modules, where each captures a different aspect of learning the behaviour of users on the macro level. All sub-modules use the attributes gathered from the macro

level, while two sub-modules use also the attributes from the micro level.

The used macro attributes are, for example, the time in the day, the day of the week (e.g. Saturday), the specific date in relation to the month (e.g. each first Monday in a month) etc. Each date also relates to a specific user in a specific way – e.g. normal working days, vacations, reported sick leaves etc. The third relation deals with previous entries in a specific time period, e.g. the last hour or on each Monday. The data features include timing of entries of the person himself or in relation to any other access of any other person. For example, appropriate input data features and examples enable finding patterns such as person #1 and person #2 always enter at the same access point inside a one minute interval.

It is important to note that learning on the macro level could be much more powerful if we had more than one access point in the system. The combination of entries at different access points gives the macro module more information and therefore facilitates the verification of users.

6.3.1 Macro tree learning

This sub-module constructs decision trees by using only attributes on the macro level. All regular entries of a person were used as positive learning examples, while all other entries were used as negative learning examples. The sub-module can classify a new access either as OK or as an alarm.

Several algorithms from machine learning toolkits Weka and Orange were tested for this task and they performed similarly. We finally chose the J48 algorithm, which is the Java implementation of Quinlan's algorithm C4.5 for constructing decision trees (Quinlan 1993). A major addition to the plain classification using decision trees was the conception of the explanation. When a classification occurs, the user interface displays on the screen the tree used for this classification. The path leading from the root to the chosen leaf is then coloured according to the classification output – either green for an OK access or red for an alarm. This gives the security personnel a comprehensive explanation of the decision of the sub-module.

6.3.2 Macro micro tree learning

This sub-module works as the previous one, with the difference that here, also the attributes from the micro level are included. Typically, the trees constructed using macro micro tree learning include some macro and some micro attributes, often gathered separately in sub-trees.

6.3.3 Macro micro LOF learning

In this sub-module, the macro and micro attributes are used by the algorithm LOF (see Subsection 6.2). While the LOF algorithm in micro learning worked only on three attributes, meaning that its results could be easily visualized, the number of attributes here is higher and other methods (such as parallel coordinates) need to be used for explanation purposes.

Classifications of all the three sub-modules are combined together using voting to represent the joint macro module classification.

6.4 Video learning

This module is essentially different from all the other ones as it learns from the camera recordings by using histograms of optical flows. More on this method can be found in (Kolbe et al. 2005; Perš et al. 2007; Sidenbladh and Black 2003).

6.5 Joint classification

In the first version, the system tackles the decisions of the four modules as equally important parts. The results are first sorted according to their value – from OK, through warning to alarm. Then, using a threshold value that is set by the security personnel that handles the system, the results of the modules are combined into the joint result in the following way. If, for example, the threshold is set to 2/4 (two modules out of four), the joint classification is equal to that of the second of the four sorted results. The higher the threshold value the stricter the system. More advanced methods are in progress (Verlinde et al. 2000; Gams 2001).

7 Empirical verification

While our implementation includes four sensors (door sensor, identity card reader, fingerprint reader and camera) and four independent modules (expert-defined rules, micro learning, macro learning and visual learning), we left out the camera and visual learning module from the presented experiments as it was not available for the tests.

Five people accessed this system – each first completing around 40 regular entries that served as learning examples for the three modules. After that, each person made another ten regular entries for testing the system. For the purpose of scientific evaluation we have performed the "fake-identity" experiment, in which each person successfully cheats the identity card reader and the fingerprint reader so that it can impersonate any other person. In this way, the testing data of one person can be used also as testing data for the other four people.

The results of this experiment are presented in Table 1. The first row represents the correct entries of the right persons. The first column represents "false" identities under which the classification system "saw" the access. There is a number of attack scenarios in which the attacker bypasses the sensory system, e.g. by a stolen identity card and a fake fingerprint. Another case would be a break in the database corrupting static data and thus faking identity, but not being able to fake the classification system output which appears directly on a screen as a result of dynamic computing.

pretender	real	#1			#2			#3			#4			#5			all		
		A	W	OK	A	W	OK	A	W	OK	A	W	OK	A	W	OK	A	W	OK
#1	rules	0	0	10	0	0	10	0	0	10	0	0	10	0	0	10	0	0	50
	micro	0	0	10	1	2	7	10	0	0	4	5	1	10	0	0	25	7	18
	macro	0	3	7	8	2	0	10	0	0	10	0	0	9	1	0	37	6	7
	together	**0**	**3**	**7**	**8**	**2**	**0**	**10**	**0**	**0**	**10**	**0**	**0**	**10**	**0**	**0**	**38**	**5**	**7**
#2	rules	0	0	10	0	0	10	0	0	10	0	0	10	0	0	10	0	0	50
	micro	0	1	9	0	0	10	5	5	0	2	0	8	6	4	0	13	10	27
	macro	10	0	0	0	1	9	1	9	0	0	10	0	10	0	0	21	20	9
	together	**10**	**0**	**0**	**0**	**1**	**9**	**5**	**5**	**0**	**2**	**8**	**0**	**10**	**0**	**0**	**27**	**14**	**9**
#3	rules	0	0	10	0	0	10	0	0	10	0	0	10	0	0	10	0	0	50
	micro	9	1	0	7	1	2	0	0	10	1	0	9	4	4	2	21	6	23
	macro	10	0	0	4	6	0	0	0	10	1	9	0	10	0	0	25	15	10
	together	**10**	**0**	**0**	**7**	**3**	**0**	**0**	**0**	**10**	**1**	**9**	**0**	**10**	**0**	**0**	**28**	**12**	**10**
#4	rules	0	0	10	0	0	10	0	0	10	0	0	10	0	0	10	0	0	50
	micro	4	4	2	0	0	10	0	0	10	0	1	9	0	1	9	4	6	40
	macro	2	7	1	0	4	6	0	0	10	0	0	10	0	0	10	2	11	37
	together	**6**	**3**	**1**	**0**	**4**	**6**	**0**	**0**	**10**	**0**	**1**	**9**	**0**	**1**	**9**	**6**	**9**	**35**
#5	rules	0	0	10	0	0	10	0	0	10	0	0	10	0	0	10	0	0	50
	micro	10	0	0	9	1	0	10	0	0	7	1	2	0	0	10	36	2	12
	macro	10	0	0	9	1	0	10	0	0	10	0	0	0	1	9	39	2	9
	together	**10**	**0**	**0**	**9**	**1**	**0**	**10**	**0**	**0**	**10**	**0**	**0**	**0**	**1**	**9**	**39**	**2**	**9**
all	rules	0	0	50	0	0	50	0	0	50	0	0	50	0	0	50	0	0	250
	micro	23	6	21	17	4	29	25	5	20	14	7	29	20	9	21	99	31	120
	macro	32	10	8	21	14	15	21	9	20	21	19	10	29	2	19	124	54	72
	together	**36**	**6**	**8**	**24**	**11**	**15**	**25**	**5**	**20**	**23**	**18**	**9**	**30**	**2**	**18**	**138**	**42**	**70**

Table 1. Evaluation of the expert-defined rules, micro learning and macro learning on 5 persons.

Whatever the case, the experiment enables to check the success rate of the stand-alone modules and the integrated system (the threshold was set to 3/3). For example if considering only micro learning, it seems obvious that some people enter in a different manner due to different physical properties, but some of them are physically and motorically similar and therefore it is possible to observe this phenomenon in Table 1. E.g. person #1 (a small woman) caused the alarm of the micro module in nine and ten out of ten cases when entering as person #3 or person #5 (both tall men). When entering as person #4 (a small man), the micro module triggered the alarm only four times, but classified the access as OK only two times. Surprisingly, person #2 (a strong middle sized man) typically entered in a similar way as person #1 (a small woman), so she was able to successfully mislead the micro module in nine out of ten cases.

As seen from Table 1, each module has its own strong and weak points. The integrated system without the visual learning classified as OK 88% of all the tested regular entries and as alarm 69% of all the irregular entries as presented in Table 2 and Table 3.

In practical experiments of several additional scenarios, the expert-defined rules and the video learning module proved quite successful on their own, and the overall performance improved as well. Due to various tests with different scenarios and under different circumstances, overall statistics of those tests are not presented through tables.

	A	W	OK
rules	0%	0%	100%
micro	0%	2%	98%
macro	0%	10%	90%
together	**0%**	**12%**	**88%**

Table 2. Statistics for regular accesses.

	A	W	OK
rules	0%	0%	100%
micro	50%	15%	36%
macro	62%	25%	14%
together	**69%**	**18%**	**13%**

Table 3. Statistics for irregular accesses (impersonation of users).

8 Conclusions

We have designed and tested an intelligent high-security access control system consisting of four sensors (door sensor, identity card reader, fingerprint reader and camera) and four independent modules (expert-defined rules, micro learning, macro learning and visual learning). The emphasis was on modifying and applying

intelligent machine learning methods to distinguish regular entries from faulty or fake ones.

The methods were tested in an experimental setting, where each of the five tested persons tried to impersonate the other four. Results from these tests are encouraging.

The experiment can also be seen as introducing intelligence into the environment. Indeed, the applied methods introduce intelligence on top of existing hardware and software solutions, improving their performance and making activities comprehensible to human users, while at the same time not burdening them. A small drawback in using this system is that it first needs to learn the regular behaviour of users, which means that it can be used only after an amount of accesses of a user have been made. Also, if a person acquires come kind of disability (for example, by braking an arm), the learning must start anew. Furthermore, several parameters need to be set in order for the system to function properly. While this can sometimes be difficult, it gives the system the necessary flexibility that enables its application for different necessities.

In summary, the machine intelligence security layer that learns from previous entries seems to be an important additional mechanism improving overall security and quality of life in modern times.

Acknowledgement

The overall project was funded partly by the Slovenian Ministry of Defence and partly by the Slovenian Research Agency.

The authors thank all the members of the research team, in particular Jana Krivec, Robert Blatnik, Erik Dovgan, Boštjan Kaluža, Aleš Tavčar.

References

[1] Albus, J. S., Meystel, A. M. (2001). *Intelligent Systems: Architecture, Design, Control*. Wiley-Interscience.

[2] Ashbourn, J. (2003). *Practical Biometrics: From Aspiration to Implementation*. Springer.

[3] Breunig, M. M. (2001). *Quality Driven Database Mining*. PhD thesis, University of Munich.

[4] Breunig, M. M., Kriegel, H.-P., Ng, R. T., Sander, J. (2000). LOF: Identifying density-based local outliers. In: *Proceedings of the International Conference on Management of Data SIGMOD'00*, pp. 93–104.

[5] Demšar, J., Zupan, B., Leban, G. (2004). *Orange: From experimental machine learning to interactive data mining*, White Paper, Faculty of Computer and Information Science, University of Ljubljana.

[6] Eskin, E., Arnold, A., Prerau, M., Portnoy, L., Stolfo, S. (2002). A geometric framework for unsupervised anomaly detection: Detecting intrusions in unlabeled data. *Data Mining for Security Applications*. Kluwer Academic Publishers.

[7] Gams, M. (2001). *Weak intelligence: Through the principle and paradox of multiple knowledge*. Advances in Computation: Theory and Practice, vol. 6. Nova Science Publishers.

[8] García Adeva, J. J., Pikatza Atxa, J. M. (2007). Intrusion detection in web applications using text mining. *Engineering Applications of Artificial Intelligence*, vol. 20, no. 4, pp. 555-566. Elsevier.

[9] Hodge, V. J., Austin, J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85-126. Springer.

[10] Hopgood, A. A. (2000). *Intelligent Systems for Engineers and Scientists*. CRC Press.

[11] Jain, L. C., Halici, U., Hazashi, I., Lee, S. B., Tsutsui, T. (1999). *Intelligent Biometric Techniques in Fingerprint and Face Recognition*. CRC Press.

[12] Kolbe, M., Gams, M. (2006) Towards an intelligent biometric system for access control. In: *Proceedings of the 9th International multiconference Information Society - IS 2006*, vol. A, pp. 118-122. Jožef Stefan Institute.

[13] Kolbe, M., Gams, M., Kovačič, S., Perš, J. (2005). A system for automatic behaviour recognition based on computer vision (in Slovene). In: *Proceedings of the 8th multiconference Information Society IS 2005*, pp. 357-361. Jožef Stefan Institute.

[14] Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, vol. 31, no. 3, pp. 249-268. Slovene Society Informatika.

[15] Lamborn, P., Williams, P. J. (2006). Data fusion on a distributed heterogeneous sensor network. In *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 6242, pp. 62420R.1-8. SPIE, Bellingham, Washington.

[16] Lane, T., Brodley, C. E. (1999). Temporal sequence learning and data reduction for anomaly detection. *ACM Transactions on Information System Security*, vol. 2, no. 3, pp. 295-331. Association for Computing Machinery.

[17] Li, C., Yang, Y.-X., Niu, X.-X. (2006) Biometric-based personal identity-authentication system and security analysis, *Journal of China Universities of Posts and Telecommunications*, vol. 13, no. 4, pp. 43-47. Elsevier.

[18] Lumini, A., Nanni, L. (2006). An advanced multi-modal method for human authentication featuring biometrics data and tokenised random numbers. *Neurocomputing*, vol. 69, no. 13-15, pp. 1706-1710. Elsevier.

[19] Markovitch, S., Rosenstein, D. (2002). Feature generation using general constructor functions. *Machine Learning*, vol. 49, no. 1, pp. 59-98. Springer.

[20] Mitchell, T. M. (1997). *Machine Learning*. McGraw Hill.

[21] Perš, J., Kristan, M., Perše, M., Kovačič, S. (2007). Motion based human identification using histograms of optical flow. In: *Proceedings of the 12th Computer Vision Winter Workshop CVWW 2007*, pp. 19-26. Graz University of Technology.

[22] Quinlan, J. R. (1993) *C4.5: Programs for Machine Learning*. Morgan Kaufmann.

[23] Revett, K., Gorunescu, F., Gorunescu, M., Ene, M., Tenreiro de Magalhaes, S., Dinis Santos, H. M. (2007). A machine learning approach to keystroke dynamics based user authentication. *International Journal of Electronic Security and Digital Forensics*, vol. 1, no. 1, pp. 55-70. Inderscience Enterprises.

[24] Sidenbladh, H., Black, M. J. (2003). Learning the statistics of people in images and video. *International Journal of Computer Vision*, vol. 54, no. 1-3, pp. 181-207. Kluwer Academic Publishers.

[25] Toledano, D. T., Pozo, R. F., Trapote, A. H., Gómez, L. H. (2006). Usability evaluation of multi-modal biometric verification systems. *Interacting with Computers*, vol. 15, no. 5, pp. 1101-1122. Elsevier.

[26] Turban, E., Aronson, J. E., Liang T.-P. (2004). *Decision Support Systems and Intelligent Systems*. Prentice Hall.

[27] Tušar, T., Gams, M. (2006). Outlier detection in an access control system (in Slovene). In *Proceedings of the 9th International multiconference Information Society - IS 2006*, vol. A, pp. 136-139. Jožef Stefan Institute.

[28] Verlinde, P., Chollet, G., Acheroy, M. (2000). Multi-modal identity verification using expert fusion. *Information Fusion*, vol. 1, no. 1, pp. 17-33. Elsevier.

[29] Wayman, J., Jain, A., Maltoni, D. Maio, D. (eds.) (2004). *Biometric Systems: Technology, Design and Performance Evaluation*. Springer.

[30] Wilson, D. L. (2005). Intelligent video systems for perimeter and secured entry access control. In *Proceedings of the 39th Annual IEEE International Carnahan Conference on Security Technology ICCST 2005*, pp. 260-262. IEEE.

[31] Witten, I. H., Frank, E. (2005). *Data Mining – Practical Machine Learning Tools and Techniques*, Chapter 4: Algorithms: The basic methods, pp. 97-105. Elsevier.

[32] Yu, L., Liu, H. (2004). Efficient Feature Selection via Analysis of Relevance and Redundancy. *Journal of Machine Learning Research*, vol. 5, pp. 1205-1224. MIT Press.

[33] Zhang, S., Zhang, C., Yang, Q. (2003). Data Preparation for Data Mining. *Applied Artificial Intelligence*, vol. 17, no. 5-6, pp. 375-381. Taylor & Francis.

A Qualitative Decision-Support Model for Evaluating Researchers

Katerina Taškova
Department of Computer Systems
Jožef Stefan Institute, Ljubljana, Slovenia
E-mail: katerina.taskova@ijs.si

Daniela Stojanova
Slovenian Forestry Institute, Ljubljana, Slovenia
E-mail: daniela.stojanova@gozdis.si

Marko Bohanec and Sašo Džeroski
Department of Knowledge Technologies
Jožef Stefan Institute, Ljubljana, Slovenia
E-mail:[marko.bohanec, saso.dzeroski]@ijs.si

Keywords: evaluation of researchers, research performance indicators, qualitative multi-attribute model, method DEXi, sensitivity analysis

Received: June 14, 2007

The evaluation of research work is an essential element of the scientific enterprise. In general, the evaluation of researchers and their work is highly dependent on the social and economic condition of the country in which the researchers work. The most commonly used form of evaluation is based on peer review. In Slovenia, a quantitative model for evaluating researchers has been developed and used by the Slovenian Research Agency, which has been criticized by the public. In order to alleviate some of the problems with this model and motivate further discussion on this issue, we propose an alternative qualitative model. The model belongs to the paradigm of hierarchical multi-attribute models and has been developed after a literature survey on existing models in foreign countries.

Povzetek: Članek prikazuje kvalitativni večparametrski model za ocenjevanje raziskovalcev in primer njegove uporabe pri ocenjevanju raziskovalcev s področja računalniških znanosti.

1 Introduction

Evaluation is an essential characteristic of human activity and is perhaps the single most important and sophisticated cognitive process in the repertoire of human reasoning and logic. It is also the one that has defied adequate explanation for nearly two millennia [1]. Without evaluation there is simply no means for distinguishing the bad from the good, the worthwhile from the worthless, and the significant from the insignificant.

In science, evaluation has been an essential element of the scientific enterprise, even before the appearance of the first scientific journals, usually in the form of peer review. In the past few decades, the evaluation of scientific research, and in particular researcher performance, has changed substantially in terms of scale and scope [2], as well as methodology. In part, these changes have occurred as a result of attempts to guide, regulate and control research agendas and priorities, not only in regard to distributing research funds, but also to influence the system of scientific research itself [3]. Therefore, criteria other than strictly scientific ones (e.g., social and political criteria) have been introduced, which further increase the complexity of evaluation. The new pillars of research evaluation include: governments; politicians; the media; social movements; and non-governmental organizations.

Typically, the evaluation takes place at a national level and each country has its own national model for evaluating research and allocating research funds. In Slovenia, such a model has been built by the Slovenian Research Agency (ARRS) [5]. This model uses data from the Slovenian Current Research Information System (SICRIS) [4] and the on-line bibliographic database COBISS [6] that maintain detailed information for every registered researcher in Slovenia. It also relies on the Web of Science database for citation information [9]. Different variants of the model are used for different disciplines and for evaluating applications to different ARRS's calls (for young researchers, for project leaders etc). Several of the variants of the model have been published together with the specific calls; these use several scientific performance indicators and combine them in a linear fashion. In the past decade, the model has been a very popular topic of discussion and criticism from the public and especially from researchers and research organizations.

Motivated by this situation, we have developed a system for evaluating researchers that uses the paradigm of qualitative multi-attribute modeling (see section 2). Our qualitative model was built on the basis of literature survey and takes into account existing foreign models for evaluation of researchers and their work. The model proposes a new methodological approach, based on qualitative multi-attribute modeling, and uses some new indicators. We propose it as an initial alternative to the existing model used by ARRS and hope to motivate further discussion on this important topic in Slovenia.

2 Methodology

Methodologically, we have taken the approach of model-based decision support [10]. We used the software tool DEXi [11] to construct a qualitative multi-attribute model aimed at evaluating and analyzing researchers. DEXi is particularly suitable for a hierarchical decomposition of evaluation problems that require judgment and qualitative reasoning.

A DEXi model is characterized by the following:

- Each model consists of a number of hierarchically structured variables called *attributes*.
- *Input attributes* are terminal nodes of the hierarchy.
- Attributes are aggregated through several levels of *aggregate attributes* into the overall assessment, which is represented by the *root attribute* of the hierarchy.
- All the attributes in the model are *qualitative*: they can take only discrete symbolic values.
- The aggregation of values in the model is defined by *decision rules*.

An example of decision rules is shown later in Fig. 3. There, each row represents a decision rule that maps two attributes, *Quality* and *Relevance*, to the *Evaluation of Researchers*. For instance, rule 1 in Fig. 3 states that if *Quality* is "Very Low" and *Relevance* is "Medium" or worse, then the *Evaluation* is "Unsatisfactory".

The model is gradually hand-crafted through four steps [10]: (1) identifying attributes, (2) structuring attributes, (3) defining attribute scales, and (4) defining decision rules. If necessary, these steps can be iterated. The model-building process is supported by the software tool DEXi, which facilitates the development of attribute trees, definition of decision rules, evaluation and analysis of options, and graphical output.

DEXi is freely available for download from http://kt.ijs.si/MarkoBohanec/dexi.html.

Usually, DEXi models are developed in collaboration between decision analysts and experts in the given field. Typically, experts suggest attributes and decision rules, while decision analysts conduct the process and define components of the model. The decision rules can be defined explicitly in tabular form or implicitly by specifying the relative importance (weight) of the contributing attributes.

The importance of attributes is most often modeled by weights in conventional multi-attribute models [12]. Each attribute is given a weight that defines the impact of that attribute to the final evaluation: the higher the weight, the more important the attribute. In DEXi, the relationship between attributes is modeled by decision rules and, in principle, there is no need for weights. However, for comparison with conventional methods, DEXi does use attribute weights; it can approximately transform decision rules to weights and vice versa:

- *From decision rules to weights* [13]: DEXi regards the currently defined decision rules as if they were points in a multi-dimensional space and approximates them by a hyperplane, using a least-squares linear regression method. From the hyperplane, it estimates approximate average weights of attributes. In Fig. 3, the weights of *Quality* and *Relevance*, obtained in this way, are 71 % and 29 %, respectively.
- *From weights to decision rules* [11]: In this case, the given weights define a multi-dimensional hyperplane, which is used to construct a complete table of decision rules. Again, each decision rule is considered to be a point in the space whose value is approximated from the hyperplane. The table constructed in this way is typically used to provide an initial ruleset, which is then reviewed by the decision-maker and possibly modified on a rule-by-rule basis.

3 Indicators of research performance

Applying research performance indicators in practice is not a straightforward task. It is important to clarify what role the indicators will play in the assessment of research, which indicators should be selected, and what possible unintended consequences could arise from their application. The problem with all quantitative indicators is that research practices vary across different fields, and it is necessary to determine what level of aggregation is to be used and the form in which the results will be presented. The vast majority of the literature discusses these issues for bibliometric indicators only. However, they affect all quantitative indicators.

A literature review was undertaken to examine quantitative performance indicators used in the evaluation of research. Quantitative evaluations of research have generally been conducted by scientometricians, bibliometricians, information and library scientists, and used indicators of quantity, quality, impact, or influence of research [7].

The indicators can be easily divided into bibliometric and non-bibliometric. *Bibliometric* indicators are based on published literature in all of its forms – journal articles, monographs, book chapters, conference papers, patents, and citations. *Non-bibliometric* measures encompass all other readily quantifiable indicators, such as the ability to attract external funding and measures of esteem: honors and

awards, editorship of journals, membership of major national and international professional societies, keynote addresses, PhD students data, etc. [8]. However, we should be careful when using non-bibliometric indicators because they can be a poor reflection of research activities in areas of applied research. Patents are regarded as a much better indicator of the output in these disciplines. Patent indicators are often used to measure the economic or innovative strength of a country in a certain area; thus, many analyses are undertaken on the macro-level in a cross-country comparison. The simplest patent indicator is the number of patents.

The number of citations is a measure of the strength of influence of a body of research, when applied to sufficiently large aggregates. Citation analyses are more difficult to undertake than publication analyses. The citations used in standard bibliometric analyses are the references contained in selected journals to other journals in the Web of Science (WoS) framework [9].

In using the citation indicator, we follow ARRS and use a citation indicator slightly different from the standard definition. Namely, due to the difficulty in obtaining detailed and consistent necessary information on citations (which is stored in different databases, using different formats of records, etc.), we use a combined approach. This approach defines the citation indicator as a weighted sum of the number of cited papers and the normalized number of citations, taking into account only plain citations (without self-citations).

We will use this citation indicator in the models together with the normalized number of citations. The former is important in situations where only one paper from a researcher's bibliography has been cited, while the rest of the work is unknown. Self-citations are excluded from the analysis.

The input to the present ARRS' model does not cover the entire space of quantitative indicators. Most notably, information, such as citations of books and citation of papers published at international conferences, are not included. Also, citation analyses are based only on the number of citations and not the number of cited papers. For some research areas (e.g. computer science), this can have a high influence on the overall evaluation.

Another important attribute is the ability to attract external funding. ARRS collects detailed information on the level of external funding attracted by individual researchers, which however is not publicly available. In our model we use the information on the number of European and National research projects as a proxy for this information.

4 Definition of the model

To define a DEXi model, we first need to identify the input attributes. We then have to specify the hierarchical structure of the model and the scales of the attributes. Finally, we need to define the decision rules for each internal node in the hierarchy. Below, we discuss each of these issues for our model for evaluating researchers.

The input attributes in our case are the performance indicators discussed in Section 3. More specifically, at the lowest level of the hierarchy we have the following attributes: *Indexed journals, Other journals, Conference publications, Monographs and other completed work, Impact, National projects* and *EU projects, SU, Prizes and awards* and *Membership*. In some of the models (M2, M2a), *Impact* is not an input attribute, but rather an aggregated one, which takes as input *Norm. num. citations* and *Num. cited Papers* (as discussed in Section 3).

The hierarchy of attributes is defined as follows (Fig. 1): At the top is the root attribute *Evaluation of Researcher*. It is decomposed into two descendants: *Quality* and *Relevance*. *Quality* aggregates *Productivity* and *Impact. Productivity* reflects the bibliometric indicators and is decomposed into *Journal publications* and *Non-journal publication. Relevance* incorporates mainly non-bibliometric indicators and is divided into *Projects* and *Other,* decomposed into *SU* (COBISS Index of professional success) and *Indicators of esteem*. Fig. 1 also gives the scales of the attributes.

The input attributes under *Relevance* have two values (Yes, No), the intermediate ones have three (Low, Medium, High) and the attribute *Relevance* has four discrete ordered values (from Low to Very High). The input attributes under *Quality* have three values (Low, Medium, High), the intermediate four (from Low to Very High). *Quality* and *Evaluation* have five ordered values each (from Very Low to Very High, and from Unsatisfactory to Excellent, respectively).

Attribute	Scale
Evaluation of Researcher	Unsatisfactory; Satisfactory; Good; Very good; *Excellent*
Quality	Very Low; Low; Medium; High; *Very High*
Productivity	Low; Medium; High; *VeryHigh*
Journal publication	Low; Medium; High; *VeryHigh*
Indexed journals	Low; Medium; *High*
Other journals	Low; Medium; *High*
Non-journal publications	Low; Medium; *High*
Conference publications	Low; Medium; *High*
Monographs and other completed work	Low; Medium; *High*
Impact	Low; Medium; *High*
Relevance	Low; Medium; High; *Very High*
Projects	Low; Medium; *High*
National projects	No; *Yes*
EU projects	No; *Yes*
Other	Low; Medium; *High*
SU	Low; Medium; *High*
Indicators of esteem	No; *Yes*
Prizes and awards	No; *Yes*
Membership	No; *Yes*

Figure1: The structure and scales of the evaluation model

The input attributes are discrete with preferentially ordered scales (values are listed in Fig. 1 from the least to the most desirable one, e.g., Unsatisfactory to Excellent). Such qualitative inputs could be obtained, for example, through a peer review process, where the quality of a researcher is assessed on a qualitative scale along each individual dimension (indicator). Another approach to obtaining qualitative input values is the discretization of continuous values (which in our case are readily available for most indicators, e.g., *Num. cited Papers*).

For each aggregate attribute, decision rules were defined so that all the combinations of the input attributes' values are mapped into values of the corresponding aggregate attribute. An example ruleset is represented in Fig. 2. Each row of the table specifies the value of the aggregate attribute for one combination of input attributes values. In this way, each row can be interpreted as an if-then rule.

	Quality	Relevance	Evaluation of Researcher
	71%	29%	
1	Very Low	<=Medium	Unsatisfactory
2	Very Low	Very High	Unsatisfactory
3	<=Low	High	Satisfactory
4	Low	<=High	Satisfactory
5	Low	Very High	Good
6	Medium	<=High	Good
7	Medium:High	Low	Good
8	Medium	Very High	Very good
9	High	Medium:High	Very good
10	>=High	Medium	Very good
11	Very High	<=Medium	Very good
12	>=High	Very High	Excellent
13	Very High	>=High	Excellent

Figure 2: The topmost decision rules

The decision rules were not specified explicitly in tabular form, but rather implicitly by specifying the weight of the input attributes. As explained in Section 2, decision rules in tabular form are derived from the weights.

In fact, we developed five variants of the model, with slight variations of the tree structure and decision rules (attribute weights). The models M1, M1k, and M1a have ten, while M2 and M2a have 11 inputs (*Impact* decomposes into *Norm. num. citations* and *Num. cited Papers*). The (global) attribute weights for the models are given in Fig. 3, where the basic attributes are in bold and the difference in the tree structure are given in italics.

In all models, the contributions of *Quality* and *Relevance* to the overall evaluation are 75 % and 25 %, respectively. In the first model, M1, the local weights of the *Impact* and *Productivity* attributes to *Quality* are 70 % and 30 %. These are approximately translated into global weights of 48 % and 27 %, respectively (summing to 75 %). The 70–30 % local weights of *Impact* and *Productivity* have been changed to 60–40 % in the model M1a and to 50–50 % in the model M2a.

The 25 % global weight of *Relevance* is divided into 17 % for *Projects* (9 % for *National* and 8 % for *EU*) and (approx.) 8 % for *Other* (of which 6 % for *SU*, 1 % for *Prizes and awards* and 1 % for *Membership*). We have 70–30 % local weight of *SU* and *Indicators of esteem* to *Other*, except in M1k, where we have equal local weights (50–50 %).

As evident from the discussion above, both local and global attribute weights are defined. The local weights always refer to a single aggregate attribute, so the sum of the weights of the immediate descendants of an aggregate attribute is 100 %. Global weights, on the other hand, take into account the structure of the evaluation model and relative importance of aggregate attributes. The sum of the global weights of all input attributes (up to rounding errors) is also equal to 100 %. At the topmost level of the hierarchy, local weights are equal to global weights.

Evaluation	Mk1	M1	M1a	M2	M2a
Quality	75	75	75	75	75
Productivity	27	27	36	27	43
Journal	17	17	23	17	27
Indexed	**14**	**14**	**18**	**14**	**22**
Other	**3**	**3**	**5**	**3**	**5**
Non-journal	10	10	14	10	17
Conference	**5**	**5**	**7**	**5**	**8**
Monographs	**5**	**5**	**7**	**5**	**8**
Impact	**48**	**48**	**39**	48	32
Norm. Num citations	-	-	-	***34***	***23***
Num. Cited papers	-	-	-	***14***	***9***
Relevance	25	25	25	25	25
Projects	17	17	17	17	17
National	**9**	**9**	**9**	**9**	**9**
EU	**9**	**8**	**8**	**8**	**8**
Other	7	7	7	7	7
SU	**5**	**6**	**6**	**6**	**6**
Indicators of esteem	3	2	2	2	2
Prizes & Awards	**1**	**1**	**1**	**1**	**1**
Membership	**1**	**1**	**1**	**1**	**1**

Figure 3: Global attribute weights in the evaluation models

5 Using and evaluating the models

In this section, we discuss how to use the developed model to evaluate researchers and illustrate its use on two sets of researchers from Slovenia. We first describe the data used, i.e., the two sets of researchers and the procedures used to discretize the input attributes. We then present a graphical description of the evaluation of four researchers. Finally, we analyze the distribution of the evaluation grades for the two sets of researchers for the several variants of the evaluation model that we have developed (the five variants, M1, M1a, M1k, M2, and M2a are described in Section 4).

5.1. Data, discretization and example use

The data we used to illustrate the use of our model(s) and analyze their behaviour concerned two batches of data on Slovenian researchers. The data were extracted by querying the COBISS database that maintains detailed data about the work of researchers in Slovenia [6]. The extracted data covered the performance of researchers in the time interval 2002–2006, which was consistent with several current ARRS's calls at the time of preparing this publication. The smaller dataset included 14 researchers, 12 of which were from the area of computer science, while the larger comprised 171 researchers, all from computer science.

The original data did not contain information on all basic attributes used in our models, hence there are undefined values for the basic attributes that are descendants of the *Indicators of esteem* attribute in all models. Actually, we used the feature of DEXi, which allows for the values of input attributes to be incompletely defined or undefined altogether. In this context completely defined means that a specific single value from the corresponding scale is given, incompletely defined means that a range of values is given instead of a specific one, and undefined means that the entire range of possible values is allowed for the attribute.

The input values of the basic attributes of the evaluated researchers (the smaller dataset is given in Fig. 4) were obtained by applying discretization to the continuous values space across which all basic attributes were initially defined. Discretization was applied to all input attributes, except for *Prizes and Awards* and *Membership*. Two different discretization approaches were taken.

The first discretization approach used a threshold. Above the threshold, the best qualitative value (High) was assigned. The interval below the threshold was divided into three equal subintervals, the first of which was also mapped to High, the remaining intervals were mapped to Medium and Low, correspondingly. As a threshold, we used the performance of the top 1 % of the 171 researchers in the first case and the top 10 % of the 14 researchers in the second case (This means the top two researchers). Ranking and discretization was performed for each indicator separately.

The second approach to discretization was based on calculating percentiles. Values belonging to the interval between the 25^{th} and 75^{th} percentile were classified as "Medium". Values below the 25^{th} percentile were classified as "Low" and values above the 75^{th} percentile as "High".

We have applied the developed model (and its variants) to the researchers from the two datasets. Let us first look at an illustration of using the basic model (M1) on four researchers from the smaller dataset. The evaluations of researchers X3, X6,Y1 and Y2 from the table in Fig 4 are depicted in graphical form in Fig 5.

The four dimensions depict the values of the attributes *Relevance, Impact, Journal publications,* and *Non-journal publications.* These are given for each of the researchers, together with the overall evaluation. We can see, for example, that X3 is evaluated as Satisfactory primarily because of the low impact. On the other hand, Y1 is evaluated as very good because of the high relevance, (very) high publications and medium impact. We can visually asses that Y1 is evaluated better than X6 as the corresponding rectangle for X6 is subsumed by the rectangle for Y1.

5.2 Distribution of the evaluation outcomes produced by the different DEXi models

After the illustrative use of the basic DEXi model on some of the researchers from the smaller set, we systematically applied the different variants of the model to both sets of researchers. The goal we had in mind was to study the similarities and differences of the different models. In particular, we compare the models in terms of the distribution of their outcomes, i.e., percentage of researchers obtaining each of the evaluation grades.

For the set of 14 researchers, we have evaluated each researcher with all the five variations of the model. By default, percentile discretization is used, but we also consider linear discretization with thresholds 1 % for M1 and M2, and 10 % for M1 (M1_Lin1, M2_Lin1, M1_Lin10). The distributions of the model outputs for each of the eight cases are depicted in the first 8 barcharts in Fig. 6.

In the case of 171 researchers, we did not consider the M2 model, due to the missing data about the number of citations per paper that this model takes as input in the calculations of the impact value. Data on *Projects* and *Membership* were also not available: we assumed completely undefined values for chart M1_D and values *Projects = Medium*, and *Membership = yes* for M1_D_PM. Finally, for the latter case, we also consider linear discretization with thresholds 1 % and 10 %.

The barcharts in Fig. 6 depict the percentage of researchers evaluated as Unsatisfactory to Excellent for each of the models and discretization approaches considered, as outlined above. The percentage evaluated as Unsatisfactory is given at the top, Excellent at the bottom of each chart.

Looking at the different charts, we can draw the following conclusions:

1. *There is a strong connection between the weights and the model output.*

   Decreasing the weight of the *Impact* attribute in favour of the *Productivity* attribute from model M1 to M1a, leads to a larger proportion of Excellent and Very Good evaluations and a smaller proportion of Good evaluations by model M1a (as compared to M1). Model M2 exhibits a similar behaviour: when we decrease the weight of attribute *Impact,* the area covered by classes Very Good and Good shrinks and the area of class Excellent extends in model M2a (as compared to M2).
2. *The model output depends strongly on the discretization applied in the pre-processing phase.*

   As mentioned above, we used two discretization techniques in order to prepare the available data in the appropriate input form for DEXi model evaluation. The barcharts M1_Lin1, M1_Lin10, M2_Lin1, M1_PM_Lin1 and M1_PM_Lin10 concern model evaluations on data obtained

Option	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Z1	Z2	V1	V2
Indexed journals	Low	Low	Medium	Medium	Medium	Medium	High	High	Medium	Medium	Medium	High	Medium	Low
Other journals	High	Low	Low	High	Low	Medium	Medium	Medium	Medium	High	Low	Medium	High	Medium
Conference publications	Medium	Low	Medium	Medium	Low	Medium	Medium	High	High	High	Medium	Medium	Low	Low
Monographs and other completed work	High	Low	High	Medium	Low	Medium	Low	Medium	Medium	Low	Medium	Medium	Low	High
Impact	Medium	Low	Low	Low	Medium	Medium	High	High	Medium	Low	High	High	Medium	Low
National projects	Medium	Medium	Medium	Medium	Medium	Medium	Medium	High	Medium	High	Medium	Medium	Medium	Medium
EU projects	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	*	*
SU	Medium	Low	High	Low	Medium	Medium	Low	High	Medium	High	Medium	Medium	Medium	Low
Prizes and awards	no	no	no	no	no	no	no	no	no	no	no	no	no	no
Membership	yes	yes	yes	yes	*	*	yes	yes	*	*	yes	yes	*	*

Figure 4: Input values of 14 evaluated researchers obtained by percentile discretization

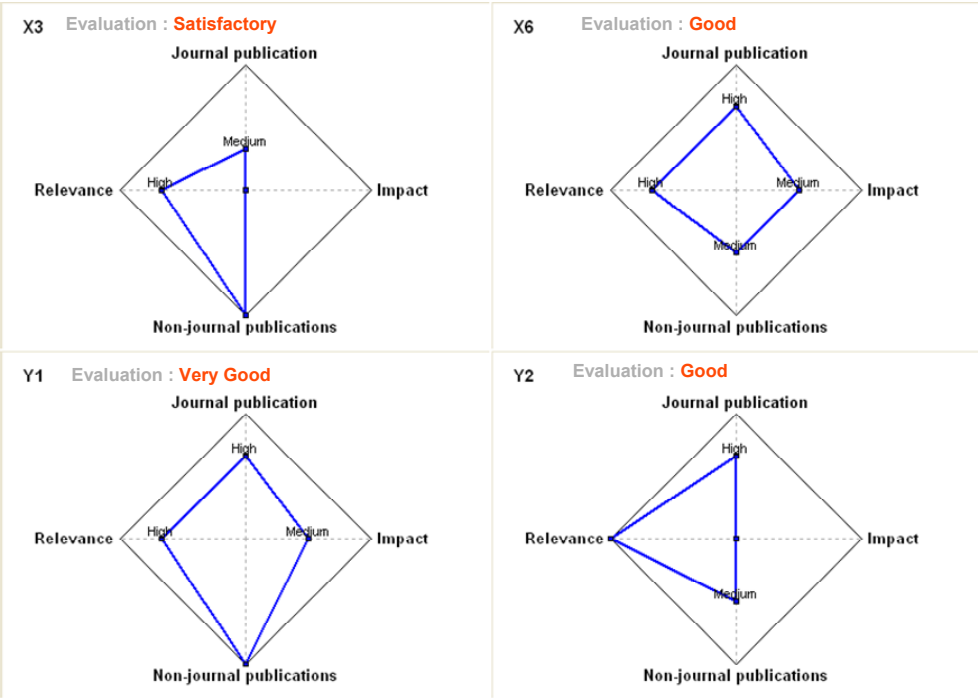


Figure 5: Chart illustration of M1 model evaluation of four researcher, using percentile based discretized data

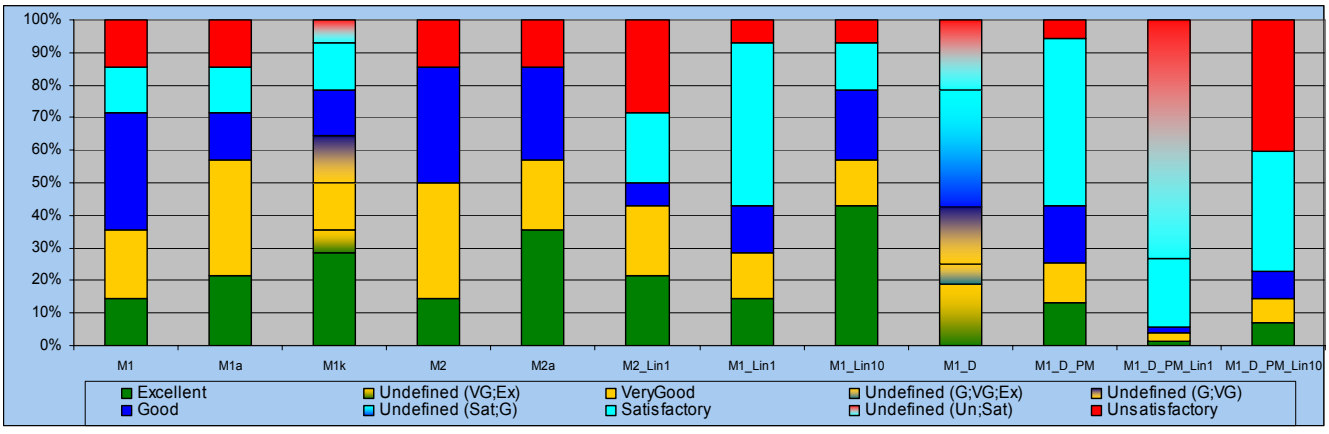


Figure 6: Distribution of evaluation results from different DEXi models, percent of classified researchers per model per class of research

by equidistant interval discretization (with thresholds 1 % and 10 % respectively). The other barcharts concern model evaluations performed on data obtained by percentile based discretizations.

Overall, percentile based discretization leads to a more balanced distribution of the researchers, with most of researchers classified in the classes Good and Very Good (M1, M1a, and M2) and Satisfactory (M1_D_PM). When equidistant interval dicretization is applied, an imbalance is visible in favour of classes Unsatisfactory (M2_Lin1, M1_D_PM_Lin10) and Excellent (M2_Lin1 and M1_Lin10).

3. *The choice of the threshold value in equidistant interval discretization can exhibit significant influence on the model behaviour.*

Using a threshold of 10 % instead of 1 % in model M1 leads to enormous shrinkage of class Satisfactory in favour of class Excellent (M1_Lin10 vs. M1_Lin1) for the small set of researchers. A similar behaviour is visible for the larger set of researchers (comparing the barcharts M1_D_PM_Lin1 and M1_D_PM_Lin10), where most of the researchers are classified as Unsatisfactory (40 %) and Satisfactory (~35 %).

The above indicates that the evaluations produced by the model are highly sensitive to the (relative) weights given to the individual attributes and to the discretization of the continuous input variables. Some of the sensitivity could be avoided by using qualitative values of the input variables directly. Such values could be derived, e.g., in a peer review process where reviewers evaluate each indicator on a qualitative scale.

The fact that the evaluations produced by the model are sensitive to the weights of attributes and to the discretization procedures also means that we can adapt the model to meet different goals without changing its structure.

Based on the overall evaluation goals and the considered field of research, we can select appropriate values for the weights and discretization approaches. For example, if we want to be strict and evaluate as excellent only a few examples of the whole population of researchers being evaluated (e.g., in the case of very limited funding), we can use model M1, or even the more strict model M1 with a linear discretization scale. On the other hand, if we want to select a larger subset for funding, we can select a model like M2 that classifies most researchers as Very Good or Good. This depends also on the set of researchers at hand, and the model can be tuned for a given set of evaluated researchers.

6 Discussion and conclusions

We have developed a hierarchical multi-attribute model for evaluating the performance of researchers and applied it to two sets of computer science researchers in Slovenia. In contrast to the current approach taken by the Slovenian Research Agency, which is quantitative and calculates a weighted sum of performance indicators, our model is qualitative and combines indicators in a sounder manner. Namely, in the case of summation we can get very high overall scores, even with very low scores along some dimensions, which is not desirable.

The model we have constructed encompasses knowledge from a wide range of studies carried out in the literature. These include researcher evaluation methods from several countries, such as the United Kingdom, the Netherlands and Australia. It is based on performance indicators that are also used in these countries.

The model that we propose can be further developed and evaluated along a number of dimensions. We have currently used weights to specify the decision rules for aggregating attributes. The intended use of weights is to provide initial rules that are reviewed and modified by a decision analyst; this was not done in our case. Manual development of decision rules is a worthwhile investment that would clearly distinguish the proposed model from quantitative linear models.

In addition, the decision support framework in which we have implemented the model has many other desirable properties. It produces evaluations for each of the intermediate levels of evaluation (such as *Quality* or *Relevance*) and provides explanations at several levels of detail. It also produces several graphical representations of the evaluations.

The proposed models are a possible alternative to the model used by ARRS and we hope it will motivate further discussion on this important topic in Slovenia.

References

[1] Scriven, M. (2001). An overview of evaluation theories. Evaluation Journal of Australasia, 1(2), 27-29.

[2] Frederiksen, L. F., Hannson, F., & Wennberg, S. B. (2003). The Agora and the role of research evaluation. Evaluation: The International Journal of Theory, Research and Practice, 9(2), 149-172.

[3] Hansson, F. (2006). Organizational use of evaluation:Governance and control in research evaluation. Evaluation: The International Journal of Theory, Research and Practice, 12 (2), 159-178.

[4] SICRIS (2007). Slovenian Current Research Information System. http://sicris.izum.si/

[5] ARRS (2007). Slovenian Research Agency. http://www.arrs.si

[6] COBISS (2007). Co-operative Online Bibliographic System & Services. http://www.cobiss.si/

[7] Coryn, C. L. S., & Hattie, J. A. (2006). The transdisciplinary model of evaluation. Journal of MultiDisciplinary Evaluation, 4, 107-114.

[8] Quantitative indicators for research assessment-a literature review

[9] Web of Science, *http://scientific.thomson.com/-products/wos/*

[10] Bohanec, M. (2003). Decision support. In: D. Mladenić, N. Lavrač, M. Bohanec, S. Moyle (Editors), Data mining and decision support: Integration and collaboration. Kluwer Academic Publishers, 23–35.

[11] Bohanec, M. (2007). DEXi: Program for Multi-Attribute Decision Making, User's Manual, Version

2.00. IJS Report DP-9596, Jožef Stefan Institute, Ljubljana, 2007.
http://www-ai.ijs.si/MarkoBohanec/dexi.html

[12] Bouyssou, D., Marchant, T., Pirlot, M., Tsoukias, A., Vincke, P. (2006). Evaluation and Decision Models with Multiple Criteria: Stepping Stones for the Analyst. Springer.

[13] Bohanec, M., Zupan, B. (2004). A function-decomposition method for development of hierarchical multi-attribute decision models. Decision Support Systems 36, 215–233.

CONTENTS OF *Informatica* Volume 31 (2007) pp. 1–489

Papers

Editorials

JOŽEF STEFAN INSTITUTE

Jožef Stefan (1835-1893) was one of the most prominent physicists of the 19th century. Born to Slovene parents, he obtained his Ph.D. at Vienna University, where he was later Director of the Physics Institute, Vice-President of the Vienna Academy of Sciences and a member of several scientific institutions in Europe. Stefan explored many areas in hydrodynamics, optics, acoustics, electricity, magnetism and the kinetic theory of gases. Among other things, he originated the law that the total radiation from a black body is proportional to the 4th power of its absolute temperature, known as the Stefan–Boltzmann law.

The Jožef Stefan Institute (JSI) is the leading independent scientific research institution in Slovenia, covering a broad spectrum of fundamental and applied research in the fields of physics, chemistry and biochemistry, electronics and information science, nuclear science technology, energy research and environmental science.

The Jožef Stefan Institute (JSI) is a research organisation for pure and applied research in the natural sciences and technology. Both are closely interconnected in research departments composed of different task teams. Emphasis in basic research is given to the development and education of young scientists, while applied research and development serve for the transfer of advanced knowledge, contributing to the development of the national economy and society in general.

At present the Institute, with a total of about 800 staff, has 600 researchers, about 250 of whom are postgraduates, nearly 400 of whom have doctorates (Ph.D.), and around 200 of whom have permanent professorships or temporary teaching assignments at the Universities.

In view of its activities and status, the JSI plays the role of a national institute, complementing the role of the universities and bridging the gap between basic science and applications.

Research at the JSI includes the following major fields: physics; chemistry; electronics, informatics and computer sciences; biochemistry; ecology; reactor technology; applied mathematics. Most of the activities are more or less closely connected to information sciences, in particular computer sciences, artificial intelligence, language and speech technologies, computer-aided design, computer architectures, biocybernetics and robotics, computer automation and control, professional electronics, digital communications and networks, and applied mathematics.

The Institute is located in Ljubljana, the capital of the independent state of S**love**nia (or S♡nia). The capital today is considered a crossroad between East, West and Mediterranean Europe, offering excellent productive capabilities and solid business opportunities, with strong international connections. Ljubljana is connected to important centers such as Prague, Budapest, Vienna, Zagreb, Milan, Rome, Monaco, Nice, Bern and Munich, all within a radius of 600 km.

From the Jožef Stefan Institute, the Technology park “Ljubljana” has been proposed as part of the national strategy for technological development to foster synergies between research and industry, to promote joint ventures between university bodies, research institutes and innovative industry, to act as an incubator for high-tech initiatives and to accelerate the development cycle of innovative products.

Part of the Institute was reorganized into several high-tech units supported by and connected within the Technology park at the Jožef Stefan Institute, established as the beginning of a regional Technology park "Ljubljana". The project was developed at a particularly historical moment, characterized by the process of state reorganisation, privatisation and private initiative. The national Technology Park is a shareholding company hosting an independent venture-capital institution.

The promoters and operational entities of the project are the Republic of Slovenia, Ministry of Higher Education, Science and Technology and the Jožef Stefan Institute. The framework of the operation also includes the University of Ljubljana, the National Institute of Chemistry, the Institute for Electronics and Vacuum Technology and the Institute for Materials and Construction Research among others. In addition, the project is supported by the Ministry of the Economy, the National Chamber of Economy and the City of Ljubljana.

Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
Tel.:+386 1 4773 900, Fax.:+386 1 251 93 85
WWW: http://www.ijs.si
E-mail: matjaz.gams@ijs.si
Public relations: Polona Strnad

INFORMATICA

AN INTERNATIONAL JOURNAL OF COMPUTING AND INFORMATICS

INVITATION, COOPERATION

Submissions and Refereeing

Please submit an email with the manuscript to one of the editors from the Editorial Board or to the Managing Editor. At least two referees outside the author's country will examine it, and they are invited to make as many remarks as possible from typing errors to global philosophical disagreements. The chosen editor will send the author the obtained reviews. If the paper is accepted, the editor will also send an email to the managing editor. The executive board will inform the author that the paper has been accepted, and the author will send the paper to the managing editor. The paper will be published within one year of receipt of email with the text in Informatica MS Word format or Informatica LaTeX format and figures in .eps format. Style and examples of papers can be obtained from http://www.informatica.si. Opinions, news, calls for conferences, calls for papers, etc. should be sent directly to the managing editor.

QUESTIONNAIRE

☐ Send Informatica free of charge

☐ Yes, we subscribe

Please, complete the order form and send it to Dr. Drago Torkar, Informatica, Institut Jožef Stefan, Jamova 39, 1000 Ljubljana, Slovenia. E-mail: drago.torkar@ijs.si

Since 1977, Informatica has been a major Slovenian scientific journal of computing and informatics, including telecommunications, automation and other related areas. In its 16th year (more than ten years ago) it became truly international, although it still remains connected to Central Europe. The basic aim of Informatica is to impose intellectual values (science, engineering) in a distributed organisation.

Informatica is a journal primarily covering the European computer science and informatics community - scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers accepted by at least two referees outside the author's country. In addition, it contains information about conferences, opinions, critical examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and information industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the Refereeing Board.

Informatica is free of charge for major scientific, educational and governmental institutions. Others should subscribe (see the last page of Informatica).

ORDER FORM – INFORMATICA

Name: .....

Title and Profession (optional): .....

.....

Home Address and Telephone (optional): .....

.....

Office Address and Telephone (optional): .....

.....

E-mail Address (optional): .....

Signature and Date: .....

Informatica WWW:

http://www.informatica.si/

Referees:

Witold Abramowicz, David Abramson, Adel Adi, Kenneth Aizawa, Suad Alagić, Mohamad Alam, Dia Ali, Alan Aliu, Richard Amoroso, John Anderson, Hans-Jurgen Appelrath, Iván Araujo, Vladimir Bajič, Michel Barbeau, Grzegorz Bartoszewicz, Catriel Beeri, Daniel Beech, Fevzi Belli, Simon Beloglavec, Sondes Bennasri, Francesco Bergadano, Istvan Berkeley, Azer Bestavros, Andraž Bežek, Balaji Bharadwaj, Ralph Bisland, Jacek Blazewicz, Laszlo Boeszoermenyi, Damjan Bojadžijev, Jeff Bone, Ivan Bratko, Pavel Brazdil, Bostjan Brumen, Jerzy Brzezinski, Marian Bubak, Davide Bugali, Troy Bull, Sabin Corneliu Buraga, Leslie Burkholder, Frada Burstein, Wojciech Buszkowski, Rajkumar Bvyya, Giacomo Cabri, Netiva Caftori, Particia Carando, Robert Cattral, Jason Ceddia, Ryszard Choras, Wojciech Cellary, Wojciech Chybowski, Andrzej Ciepielewski, Vic Ciesielski, Mel Ó Cinnéide, David Cliff, Maria Cobb, Jean-Pierre Corriveau, Travis Craig, Noel Craske, Matthew Crocker, Tadeusz Czachorski, Milan Češka, Honghua Dai, Bart de Decker, Deborah Dent, Andrej Dobnikar, Sait Dogru, Peter Dolog, Georg Dorfner, Ludoslaw Drelichowski, Matija Drobnič, Maciej Drozdowski, Marek Druzdzel, Marjan Družovec, Jozo Dujmović, Pavol Ďuriš, Amnon Eden, Johann Eder, Hesham El-Rewini, Darrell Ferguson, Warren Fergusson, David Flater, Pierre Flener, Wojciech Fliegner, Vladimir A. Fomichov, Terrence Forgarty, Hans Fraaije, Stan Franklin, Violetta Galant, Hugo de Garis, Eugeniusz Gatnar, Grant Gayed, James Geller, Michael Georgiopolus, Michael Gertz, Jan Goliński, Janusz Gorski, Georg Gottlob, David Green, Herbert Groiss, Jozsef Gyorkos, Marten Haglind, Abdelwahab Hamou-Lhadj, Inman Harvey, Jaak Henno, Marjan Hericko, Henry Hexmoor, Elke Hochmueller, Jack Hodges, John-Paul Hosom, Doug Howe, Rod Howell, Tomáš Hruška, Don Huch, Simone Fischer-Huebner, Zbigniew Huzar, Alexey Ippa, Hannu Jaakkola, Sushil Jajodia, Ryszard Jakubowski, Piotr Jedrzejowicz, A. Milton Jenkins, Eric Johnson, Polina Jordanova, Djani Juričič, Marko Juvancic, Sabhash Kak, Li-Shan Kang, Ivan Kapustøk, Orlando Karam, Roland Kaschek, Jacek Kierzenka, Jan Kniat, Stavros Kokkotos, Fabio Kon, Kevin Korb, Gilad Koren, Andrej Krajnc, Henryk Krawczyk, Ben Kroese, Zbyszko Krolikowski, Benjamin Kuipers, Matjaž Kukar, Aarre Laakso, Sofiane Labidi, Les Labuschagne, Ivan Lah, Phil Laplante, Bud Lawson, Herbert Leitold, Ulrike Leopold-Wildburger, Timothy C. Lethbridge, Joseph Y-T. Leung, Barry Levine, Xuefeng Li, Alexander Linkevich, Raymond Lister, Doug Locke, Peter Lockeman, Vincenzo Loia, Matija Lokar, Jason Lowder, Kim Teng Lua, Ann Macintosh, Bernardo Magnini, Andrzej Małachowski, Peter Marcer, Andrzej Marciniak, Witold Marciszewski, Vladimir Marik, Jacek Martinek, Tomasz Maruszewski, Florian Matthes, Daniel Memmi, Timothy Menzies, Dieter Merkl, Zbigniew Michalewicz, Armin R. Mikler, Gautam Mitra, Roland Mittermeir, Madhav Moganti, Reinhard Moller, Tadeusz Morzy, Daniel Mossé, John Mueller, Jari Multisilta, Hari Narayanan, Jerzy Nawrocki, Rance Necaise, Elzbieta Niedzielska, Marian Niedq'zwiedziński, Jaroslav Nieplocha, Oscar Nierstrasz, Roumen Nikolov, Mark Nissen, Jerzy Nogieć, Stefano Nolfi, Franc Novak, Antoni Nowakowski, Adam Nowicki, Tadeusz Nowicki, Daniel Olejar, Hubert Österle, Wojciech Olejniczak, Jerzy Olszewski, Cherry Owen, Mieczyslaw Owoc, Tadeusz Pankowski, Jens Penberg, William C. Perkins, Warren Persons, Mitja Peruš, Fred Petry, Stephen Pike, Niki Pissinou, Aleksander Pivk, Ullin Place, Peter Planinšec, Gabika Polčicová, Gustav Pomberger, James Pomykalski, Tomas E. Potok, Dimithu Prasanna, Gary Preckshot, Dejan Rakovič, Cveta Razdevšek Pučko, Ke Qiu, Michael Quinn, Gerald Quirchmayer, Vojislav D. Radonjic, Luc de Raedt, Ewaryst Rafajlowicz, Sita Ramakrishnan, Kai Rannenberg, Wolf Rauch, Peter Rechenberg, Felix Redmill, James Edward Ries, David Robertson, Marko Robnik, Colette Rolland, Wilhelm Rossak, Ingrid Russel, A.S.M. Sajeev, Kimmo Salmenjoki, Pierangela Samarati, Bo Sanden, P. G. Sarang, Vivek Sarin, Iztok Savnik, Ichiro Satoh, Walter Schempp, Wolfgang Schreiner, Guenter Schmidt, Heinz Schmidt, Dennis Sewer, Zhongzhi Shi, Mária Smolárová, Carine Souveyet, William Spears, Hartmut Stadtler, Stanislaw Stanek, Olivero Stock, Janusz Stokłosa, Przemysław Stpiczyński, Andrej Stritar, Maciej Stroinski, Leon Strous, Ron Sun, Tomasz Szmuc, Zdzislaw Szyjewski, Jure Šilc, Metod Škarja, Jiři Šlechta, Chew Lim Tan, Zahir Tari, Jurij Tasič, Gheorge Tecuci, Piotr Teczynski, Stephanie Teufel, Ken Tindell, A Min Tjoa, Drago Torkar, Vladimir Tosic, Wieslaw Traczyk, Denis Trček, Roman Trobec, Marek Tudruj, Andrej Ule, Amjad Umar, Andrzej Urbanski, Marko Uršič, Tadeusz Usowicz, Romana Vajde Horvat, Elisabeth Valentine, Kanonkluk Vanapipat, Alexander P. Vazhenin, Jan Verschuren, Zygmunt Vetulani, Olivier de Vel, Didier Vojtisek, Valentino Vranić, Jozef Vyskoc, Eugene Wallingford, Matthew Warren, John Weckert, Michael Weiss, Tatjana Welzer, Lee White, Gerhard Widmer, Stefan Wrobel, Stanislaw Wrycza, Tatyana Yakhno, Janusz Zalewski, Damir Zazula, Yanchun Zhang, Ales Zivkovic, Zonling Zhou, Robert Zorc, Anton P. Železnikar

Informatica

An International Journal of Computing and Informatics

Web edition of Informatica may be accessed at: http://www.informatica.si.

Subscription Information Informatica (ISSN 0350-5596) is published four times a year in Spring, Summer, Autumn, and Winter (4 issues per year) by the Slovene Society Informatika, Vožarski pot 12, 1000 Ljubljana, Slovenia.
The subscription rate for 2007 (Volume 31) is
– 60 EUR for institutions,
– 30 EUR for individuals, and
– 15 EUR for students
Claims for missing issues will be honored free of charge within six months after the publication date of the issue.

Typesetting: Borut Žnidar.
Printing: Dikplast Kregar Ivan s.p., Kotna ulica 5, 3000 Celje.

Orders may be placed by email (drago.torkar@ijs.si), telephone (+386 1 477 3900) or fax (+386 1 251 93 85). The payment should be made to our bank account no.: 02083-0013014662 at NLB d.d., 1520 Ljubljana, Trg republike 2, Slovenija, IBAN no.: SI56020830013014662, SWIFT Code: LJBASI2X.

Informatica is published by Slovene Society Informatika (president Niko Schlamberger) in cooperation with the following societies (and contact persons):
Robotics Society of Slovenia (Jadran Lenarčič)
Slovene Society for Pattern Recognition (Franjo Pernuš)
Slovenian Artificial Intelligence Society; Cognitive Science Society (Matjaž Gams)
Slovenian Society of Mathematicians, Physicists and Astronomers (Bojan Mohar)
Automatic Control Society of Slovenia (Borut Zupančič)
Slovenian Association of Technical and Natural Sciences / Engineering Academy of Slovenia (Igor Grabec)
ACM Slovenia (Dunja Mladenič)

Informatica is surveyed by: Citeseer, COBISS, Compendex, Computer & Information Systems Abstracts, Computer Database, Computer Science Index, Current Mathematical Publications, DBLP Computer Science Bibliography, Directory of Open Access Journals, InfoTrac OneFile, Inspec, Linguistic and Language Behaviour Abstracts, Mathematical Reviews, MatSciNet, MatSci on SilverPlatter, Scopus, Zentralblatt Math

The issuing of the Informatica journal is financially supported by the Ministry of Higher Education, Science and Technology, Trg OF 13, 1000 Ljubljana, Slovenia.

Volume 31 Number 4 December 2007 **ISSN 0350-5596**

Informatica

An International Journal of Computing and Informatics

Informatica **31** (2007) Number 4, pp. 355–489