ICDS 2012

The Sixth International Conference on Digital Society

ISBN: 978-1-61208-176-2

January 30- February 4, 2012

Valencia, Spain

ICDS 2012 Editors

Jaime Lloret Mauri, Polytechnic University of Valencia, Spain

Gregorio Martinez, University of Murcia, Spain

Lasse Berntzen, Vestfold University College - Tønsberg, Norway

Åsa Smedberg, Stockholm University/The Royal Institute of Technology, Sweden

ICDS 2012

Forward

The sixth edition of The International Conference on Digital Society (ICDS 2012) was held in Valencia, Spain, on January 30th – February 4th, 2012.

Nowadays, most of the economic activities and business models are driven by the unprecedented evolution of theories and technologies. The impregnation of these achievements into our society is present everywhere, and it is only question of user education and business models optimization towards a digital society.

Progress in cognitive science, knowledge acquisition, representation, and processing helped to deal with imprecise, uncertain or incomplete information. Management of geographical and temporal information becomes a challenge, in terms of volume, speed, semantic, decision, and delivery.

Information technologies allow optimization in searching an interpreting data, yet special constraints imposed by the digital society require on-demand, ethics, and legal aspects, as well as user privacy and safety.

The event was very competitive in its selection process and very well perceived by the international scientific and industrial communities. As such, it is attracting excellent contributions and active participation from all over the world. We were very pleased to receive a large amount of top quality contributions.

The accepted papers covered a large spectrum of topics related to advanced networking, applications, social networking, and systems technologies in a digital society. We believe that the ICDS 2012 contributions offered a large panel of solutions to key problems in all areas of digital needs of today's society.

We take here the opportunity to warmly thank all the members of the ICDS 2012 technical program committee as well as the numerous reviewers. The creation of such a broad and high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and efforts to contribute to the ICDS 2012. We truly believe that thanks to all these efforts, the final conference program consists of top quality contributions.

This event could also not have been a reality without the support of many individuals, organizations and sponsors. In addition, we also gratefully thank the members of the ICDS

20102 organizing committee for their help in handling the logistics and for their work that is making this professional meeting a success.

We hope the ICDS 2012 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress on the topics of the conference.

We also hope the attendees enjoyed the beautiful surroundings of Valencia, Spain.

ICDS 2012 Chairs

ICDS 2012 General Chair

Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Gregorio Martinez, University of Murcia, Spain

ICDS 2012 Advisory Committee

Lasse Berntzen, Vestfold University College - Tønsberg, Norway
Åsa Smedberg, DSV, Stockholm University/KTH, Sweden
Freimut Bodendorf, University of Erlangen, Germany
Adolfo Villafiorita, Fondazione Bruno Kessler, Italy
A.V. Senthil Kumar, Hindusthan College of Arts and Science, India
Charalampos Konstantopoulos, University of Piraeus, Greece

ICDS 2012

Committee

ICDS 2012 General Chair

Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Gregorio Martinez, University of Murcia, Spain

ICDS 2012 Advisory Committee

Lasse Berntzen, Vestfold University College - Tønsberg, Norway
Åsa Smedberg, DSV, Stockholm University/KTH, Sweden
Freimut Bodendorf, University of Erlangen, Germany
Adolfo Villafiorita, Fondazione Bruno Kessler, Italy
A.V. Senthil Kumar, Hindusthan College of Arts and Science, India
Charalampos Konstantopoulos, University of Piraeus, Greece

ICDS 2012 Technical Program Committee

Gil Ad Ariely, California State University (CSU), USA / Interdisciplinary Center(IDC) Herzliya, Israel
Adolfo Albaladejo Blázquez, Universidad de Alicante, Spain
Salvador Alcaraz Carrasco, Universidad Miguel Hernández, Spain
Shadi Aljawarneh, Isra University - Amman, Jordan
Giner Alor Hernández, Instituto Tecnológico de Orizaba-Veracruz, México
Aini Aman, Universiti Kebangsaan Malaysia, Malaysia
Pasquale Ardimento, University of Bari, Italy
Marcelo E. Atenas, Universidad Politecnica de Valencia, Spain
Charles K. Ayo, Covenant University, Nigeria
Gilbert Babin, HEC Montréal, Canada
Kambiz Badie, Iran Telecom Research Center & University of Tehran, Iran
Lasse Berntzen, Vestfold University College - Tønsberg, Norway
Aljoša Jerman Blažič, SETCCE - Ljubljana, Slovenia
Marco Block-Berlitz, Mediadesign Hochschule- Berlin, Germany
Nicola Boffoli, University of Bari, Italy
Mahmoud Boufaida, Mentouri University of Constantine, Algeria
Mahmoud Brahimi, University of Msila, Algeria
Diana Bri, Universidad Politecnica de Valencia, Spain
Luis M. Camarinha-Matos, New University of Lisbon, Portugal
Vlatko Ceric, University of Zagreb, Croatia
Walter Castelnovo, University of Insubria, Italy
Yul Chu, University of Texas Pan American, USA
David Day, Sheffield Hallam University, UK
Gert-Jan de Vreede, University of Nebraska at Omaha, USA
Prokopios Drogkaris, University of the Aegean - Karlovasi, Greece

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Polish E-Government at Local Level

Heavy Road to Citizens' Empowerment

Leszek Porebski
Department of Political Science and Contemporary History
AGH University of Science and Technology
Krakow, Poland
leszekpo@agh.edu.pl

***Abstract*—The paper evaluates the development of e-government in Polish local governments, within the framework of the role played by an individual in political processes. Presented here are the results of the empirical research carried out in the period of 2005-2009. The study comprised the assessment of the official websites of Polish counties, the secondary level of local government system. Sites of 314 counties were analyzed, with the application of the quantitative method based on Website Attribute Evaluation System. The change of citizens' position with respect to public institutions was assessed against the background of the four basic functions performed by local governments websites. They are: information, promotion, consultation and service delivery. Research results indicate that local level of Polish e-government is on the preliminary stage of development and the impact of new technologies on the model of local democracy is limited.**

Keywords-e-government; e-democracy; websites content; local democracy; local government.

I. INTRODUCTION

The use of Information and Communication Technologies (ICT) by public institutions remains one of the most popular issues undertaken by scholars dealing with social implications of the information revolution. Nevertheless, the local level still attracts much less attention than activity of parliaments, governments or governmental agencies. While the vast majority of both individual research projects and international benchmarking studies focus mostly on consequences of ICT use in the macro-scale of political and administrative processes [1], [2], [3], from the perspective of a democratic theory it is the study of the local democracy that offers an excellent insight into a political transformation. The survey of what occurs in local communities can be the preliminary stage of pointing the overall direction democracy heads for.

The present paper presents an assessment of the role played by local government websites in redefining the model of contemporary democracy. The focus of the analysis is on the position taken by the individual in his relations with the state, represented by public institutions. Without a doubt ICT modify patterns of interactions between various political actors. Nonetheless, consequences of the emergence of new technologies on the status of the citizen in relation to the state have been vigorously debated for several years. ICT enthusiasts note the definite positive the impact of technologies on political life, such as the empowerment of an individual in the realm of political communication, political participation as well as decision-making [4], [5]. More skeptical observers however point out that cyberspace in fact mirrors the real life "politics as usual" game, with the same actors dominating the scene [6]. Others claim that it is too early to prejudge the ultimate effect of ICT use in political processes [7],[8].

Is the role of the individual, especially in local democracy, enhanced by new technologies? To what extent websites of local government institutions stimulate civic activism and participatory attitudes? What type of democracy is formed by the way local authorities use the ICT? These issues are addressed in the paper, on the base of the empirical assessment of the content of Polish local government websites.

II. BASIC TERMS

E-government is the notion which is both commonly used and at the same time lacking agreed, precise meaning. It can be defined as the use of technology in the management and delivery of public services [9], or the employment of ICT to provide electronic services to citizens, businesses and organizations [10]. The same term is however described at times in much broader perspective. According to Carbo and Williams [11] the role of e-government is also to involve citizens in the democratic process and decision making in the convenient, customer-oriented and cost-effective way. Consequently, e-government cannot be reduced to the process of electronic services distribution. It is much more than merely the technological phenomenon. The essence of e-government is the reconstruction of mutual interactions between citizens and service providers [12].

In this paper, the broad concept of e-government is assumed. It comprises a few basic dimensions. Most important of them are: delivery of public services, provision of information, strengthening the public debate as well as the stimulation of citizens participation and their involvement in the decision making process. Such a wide-ranging approach to e-government makes it very close to the notion of e-

 ISBN: 978-1-61208-176-2

democracy. In fact, e-government can be considered to be the aspect of e-democracy associated with the pursuit of various types of public institutions. It obviously implies that performance of e-government considerably determines whether the overall goals of the ICT use in politics can be accomplished, regardless of how they are articulated.

III. SUBJECT AND SCOPE OF THE RESEARCH

The research project presented in this paper was dedicated to the analysis of official websites of Polish counties. In the three-tier system of local government in Poland (introduced in January 1999) municipalities are the primary units, counties are units on the secondary level and provinces make up the third tier of the system. There are 2478 municipalities, 379 counties and 16 provinces altogether.

There are two different types of counties in Poland: urban and territorial ones. Cities with population over one hundred thousand residents establish 65 urban counties. The territory of the county is in this case limited to the area of the single city – the county seat. Nevertheless, legally these cities are endowed with rights of counties. The second category of counties – territorial counties, are composed of several rural and urban municipalities. The largest city in the area is usually the seat of the county. It performs the role of educational, economic and cultural center of the region as well. Regions represented by territorial counties are very often linked with strong ties, which are rooted in shared history and common traditions. The sense of local identity is thus often preserved by both residents of the county and its local government authorities.

The elected organ of the county is the council, while the executive branch is represented by the county board. The chair of the board is in charge of both temporary works of the county administration and the execution of the policy assumed by the council. There are several statutory tasks of the county. The most important of them include: health care, social welfare, public transport and public roads maintenance, culture and tourism, education and building supervision.

In the reviewed research, websites of territorial counties were the only ones to be selected for the analysis. Urban counties were excluded from the study to ensure the internal cohesion of the sample and to allow for generalized conclusions. Cities with populations close to or even greater than half a million inhabitants are very much different from the majority of territorial counties. The latter are typically rural and sparsely populated units, often – as mentioned above – founded around common history and enduring social ties. Therefore, the assumption that both types of counties are equal (and including them within the same sample) would distort the results of the study.

Consequently, websites of all the 314 territorial counties in Poland were analyzed within the framework of the project. There were the official websites only, these maintained formally by the county office. The research was carried out for five years, from 2005 to 2009, between April and May of each year.

IV. METHOD OF THE RESEARCH

The major goal of the research project presented in this paper was the comprehensive assessment of the content of the counties' websites. The questionnaire constructed for the study was a quantitative one. It was based on the overall idea proposed by the Cyberspace Policy Research Group, known as the Website Attribute Evaluation System (WAES). The WAES is used both in the analysis of websites [13] and as the point of reference for researchers of websites performance [14], [15]. The WAES is the binary tool. It analyses the content of the website in the context of specific detailed criteria (types of information, services, web tools). The component in the content either exists or is absent. As a result, a score of either "0" or "1" is assigned to the specific criterion.

The questionnaire applied in the analysis of websites of Polish counties was founded on the same principle. In the 2005 edition of the research it included 55 detailed criteria. After minor modifications introduced in 2006 (a few criteria were substituted with new ones) the number of criteria was reduced to 54. This final version of the questionnaire was used in the research conducted from 2006 to 2009.

Prior to the beginning of the actual research, in November 2004, the preliminary, qualitative survey of several local government websites was performed. Four major aspects of website content were identified on the base of its results. They are major functions performed by websites in everyday activity of local government institutions. The functions are: information, promotion, consultation and service delivery. In the questionnaire several specific criteria were assigned to each function.

The information function is associated with the access of Internet users to various types of data. Local governments publish both basic personal information (composition of the county council and the board) and information pertaining to their work (office hours, announcements), as well as other information helpful for customers of the county administration. They include: the division of powers among various departments of the county administration and information on handling specific matters. A website of local government can be also regarded as a hub in the network connecting many different kinds of public institutions, local civic initiatives, NGOs, etc. The simplest way to facilitate this process is to place links to such organizations on the website. Accessibility of such links is also a part of information function.

Promotion is the only function performed by the website aimed mainly at non-residents of the county. That aspect of online presence includes the presentation of touristic and cultural qualities of the region (directed to individual visitors) as well as commercial assets (e.g., offers to potential investors). An important dimension of promoting the region is also the availability of website content in foreign languages.

Consultation is the most directly "political" dimension of local government websites content. It includes services and tools that stimulate public debate on local issues and enhance communication with citizens as well as civic participation.

 ISBN: 978-1-61208-176-2

Detailed criteria include the availability of email addresses to local government representatives, online polls, discussion forum or chat.

The last function, electronic delivery of public services, can be regarded as synonymous to the narrowly defined concept of e-government. It refers to interactions between local government administration and the individual (considered as a beneficiary of various services). The questionnaire applied in the research has not assessed the electronic availability of specific services. Instead, stages of online sophistication were measured. They include: downloading forms, the ability to apply online, online transactions with the office as well as the possibility of tracking the individual matter handling.

Beside the survey of four major functions the questionnaire included also a few criteria assessing the availability of additional services. They were: accessibility of the web site for persons with disabilities and presence of various types of multimedia content (pictures, audio and video materials).

All major functions performed by local government websites can be regarded as founding elements of the broadly defined domain of e-government. Therefore, the research of Polish counties websites was in fact an attempt to evaluate the standing of local dimension of e-government development during the first decade of the 21st century. It was also the indirect indicator of the current status of e-democracy in local communities.

V. THE CONTENT OF LOCAL GOVERNMENT WEBSITES. RESEARCH RESULTS

A. Overall results and the distribution of scores

As mentioned above, the maximum score that could be obtained in the research was 55 points in 2005 and 54 points in 2006-2009 period. Fig. 1 presents the average scores achieved by county websites in the consecutive years of the analysis. The results indicate that except for 2006, when slight decline in the total score was noticed, we can observe gradual and steady growth of overall sophistication of the websites. The greatest progress appeared between 2006 and 2007, followed by the decline of the pace of growth..

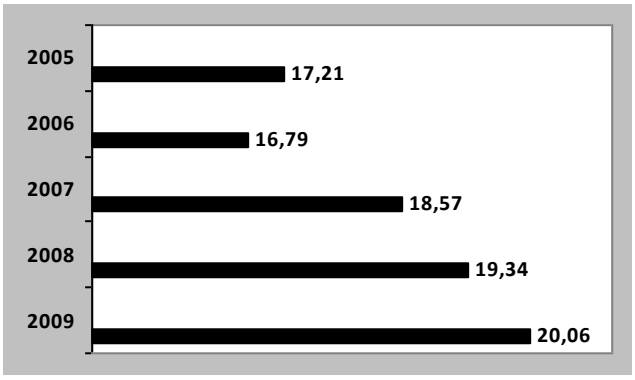


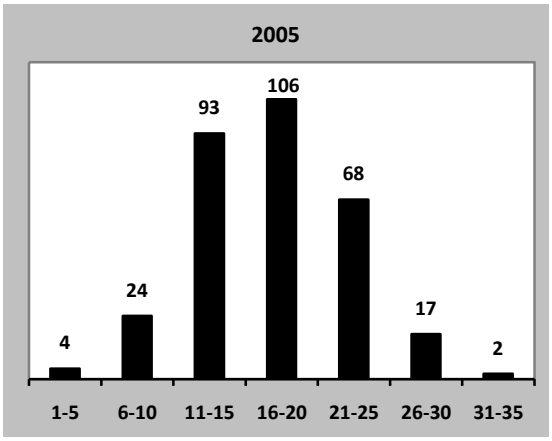
Figure 1. Average scores of local government websites.

This development can be considered disappointing, especially if compared with the rapid growth of overall ICT accessibility during the same period all over the world, including Poland. Local governments improved their web offer although they have hardly kept abreast of the overall progress in technology.

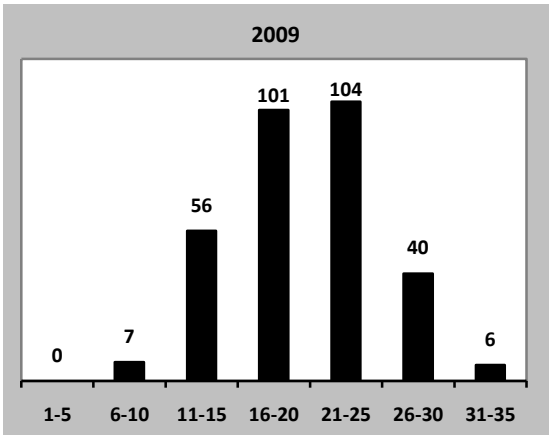
Fig. 2 depicts the distribution of scores in the first (2005) and the last (2009) year of the research project execution. In 2005 about one third of all the websites (33,7%) scored between 16 and 20 points. More than one hundred counties can be found in this interval, making it the most numerous category of scores.

Quite similar frequency of scores can be found however in the interval between 11 and 15 points, which is a category of explicitly minor scores, compared with the mean value. Altogether, as much as 85% of all the websites scored between 11 and 25 points. The single minimum result was 3 points, while the maximum score was 31 points (obtained by two websites). That proves that the gap between leaders and laggards was considerable.

Five years later, in 2009, the distribution of scores clearly leans towards results surpassing the mean value. At this time the most common category of scores is the interval between 21 and 25 points (more than both the mean and the median value).



Mean = 17,21, Median = 17, Std. Dev. = 5,20, N=314



Mean = 20,06, Median = 20, Std. Dev. = 5,08, N=314

Figure 2. Distribution of local government websites scores.

 ISBN: 978-1-61208-176-2

Simultaneously, the frequency of scores between 11 and 15 points is almost two times lower, than in 2005. In 2009 both the minimum and the maximum results have increased. They were: 7 and 34 points, respectively. Thus the whole sample remained very much differentiated, with the variation between the best and the worst websites similar to that of 2005.

B. Basic functions performance

The adequate measure of the advancement in the use of Internet capabilities is the presentation of websites score as the percentage of the total number of points, which could have been obtained. These data, concerning both the total score and specific functions, are presented in Table I.

The starting point of the project – 2005 – is the moment when less than half of criteria taken into account in the research questionnaire were fulfilled. It refers both to the overall score and to each particular function. Performance of information and promotion is much more advanced than in case of consultation and service delivery. Nevertheless, even with respect to the provision of information, the leading aspect of web use, the score is slightly below fifty percent. In 2005, electronic service delivery was especially at the preliminary stage of development. Its score strongly lagged behind all other functions, with result two times lower than the total score.

Successive years bring about the improvement of scores but the sequence of performance of specific functions continues to be stable. The most advanced dimension of websites content is the access to information, followed by promotion and consultation. Service delivery all the time remains to be the least developed. In 2009 the total score was about one fifth greater as compared to the first year of the research. Approximately the same level of increase can be observed with respect to information provision. The lowest increase occurred in case of the promotion. In fact it was the only function that has barely grown during five years of the research. Aspects of websites content which registered the most significant progress are: consultation and, in particular – service delivery. In case of the latter the spectacular growth of performance took place during the last two years of the analysis. In 2007 the score of this function was lower than in 2005, while during next two years it recorded the improvement by almost seventy percent.

Data comparison between 2005 and 2009 proves gradual transformation of the pattern of websites use by Polish local government.

TABLE I. LOCAL GOVERNMENT WEBSITES SCORES (SELECTED YEARS)

	2005	**2007**	**2009**	**2005 = 100**	
				2007	***2009***
Total score	40,00	44,21	47,76	110,5	119,4
Information	49,30	55,80	58,15	113,2	118,0
Promotion	46,86	47,14	48,00	100,6	102,4
Consultation	26,56	32,29	34,29	121,6	129,1
Service Delivery	19,33	17,00	28,75	87,9	148,7

In the first year of the research passive and one-way forms of web communication (information and promotion) were dominating in the offer of self-government institutions for the Internet users. In the successive years, mostly interactive features of websites were improved, making the content of websites more balanced and open to more active participation of citizens. It represents the typical model of e-government and e-democracy development [16].

Nevertheless, functions which are essential from the perspective of democratic theory – consultation and service delivery – are still evidently delayed. Unless the pace of described changes accelerates, websites of Polish local governments will continue to function as the electronic bulletin boards as opposed to the tools of real political and civic interaction.

C. Availability of resources encouraging the development of e-democracy

In the framework of the analysis of Polish local governments websites three functions play critical role in terms of supporting the citizens-oriented model of e-government. They are: information, consultation and service delivery. Accessibility of particular types of information, web tools and services is then the adequate measure of the "democratic maturity" of the assessed websites.

Table II presents data on selected criteria of the information function. As mentioned in the previous part of this text the provision of information is performed at relatively satisfactory level.

Basic data on the county office are available on virtually every one of the analyzed websites (however, in 2009 precise instructions on handling particular matters could have been found only on every second – 49,4% – of assessed sites). Moreover, local government site gradually becomes an information center for the local community. In the consecutive years of research there was a visible growth in the availability of resources useful in everyday life of residents. It refers to local newsletter (with information on program of movies, cultural events, etc.), links to websites of municipalities in the county as well as sites of local NGOs. In 2009 these data were available on the vast majority of the surveyed websites.

A completely different picture emerges when consultation resources are analyzed (see Table III). These are services dedicated directly to the encouragement of civic activity and public debate. Therefore, their presence on local government sites is the indicator of local authorities' readiness to face the real e-participation. Research results suggest that representatives of Polish local governments are not exactly enthusiastic about this prospect.

TABLE II. LOCAL GOVERNMENT WEBSITES SCORES (SELECTED YEARS)

Content	**2005**	**2007**	**2009**
Organization of the office	89,2	89,2	92,0
Local newsletter	54,8	58,0	70,4
Links to government websites	33,1	33,8	29,3
Links to websites of municipalities	74,2	87,6	87,6
Links to websites of local NGOs	51,6	57,9	63,7

 ISBN: 978-1-61208-176-2

TABLE III. AVAILABILITY OF SELECTED CONSULTATION RESOURCES

Content	2005	2007	2009
Email address to the office	90,8	95,5	95,5
Online discussion forum	27,1	16,2	13,7
Online poll on local issues	15,6	19,1	18,8
Chat with county officials	4,5	2,2	2,6
Interactive service for the direct contact with county officials	8,0	14,3	16,9

The only resource commonly available on the assessed websites is the office email address (it is worth stressing however that even in 2009 a few counties have not published their own email addresses on their websites). Nevertheless, making email address accessible is not the same as responsiveness. In 2009, only one fifth of websites (21,7%) replied to electronic messages sent to local authorities during the research. It proves that email is still not perceived as the regular means of communication with citizens.

All the other, much more sophisticated, resources can be only found on the relatively small part of the analyzed sites. It refers to discussion forums, online polls, chats as well as various types of interactive services facilitating contacts with representatives of local authorities. In 2009 none of these resources was available on more than one fifth of evaluated sites. Specifically, chat can be hardly found in general (only 2,6% of sites enabled the use of that service). Out of tools stimulating the public debate, in 2005 a discussion forum was the most popular one. Every fourth of assessed sites (27,1%) provided the opportunity to debate on local issues. The following years brought about the remarkable decline of the accessibility of that service (2009 – 13,7%). This tendency goes together with growing popularity of online tools, which assist in the direct contact with county officials. Both services represent however various forms of online communication. Discussion forums are open for multilateral communication, protect anonymity and provide the arena for real deliberation of public problems. Conversely, services facilitating contact with local representatives enable only bilateral interaction, during which an individual Internet user can ask a question or present his or her views to the particular county official. In addition, in 2009 more than every second of the latter services (54,7%) required revealing personal data of a user. Growing popularity of this mode of local debate proves an obvious intention of local authorities to manage the course of online dialogue with citizens. What can be observed is then the emergence of the supervised model of e-democracy.

Distribution of public services is another aspect of online pursuit, which is of fundamental importance in the context of e-government development. With regard to Polish local governments, electronic delivery of services – as mentioned before – performs at a disappointingly low level (see Table IV). A relatively well accessible option is the ability to download various forms from the county site. Close to two thirds of local governments (59,6%) provided this opportunity in 2009. On the contrary, online transactions are still hardly possible on the analyzed sites. Only seven out of 314 counties (2,2%) made that service available.

TABLE IV. AVAILABILITY OF SELECTED SERVICE DELIVERY RESOURCES

Content	2005	2007	2009
Downloading forms	46,2	55,1	59,6
Ability to apply online	2,9	2,6	28,7
Online transactions[a]	-	0,3	2,2

a. The criterion was included to the questionnaire in 2006.

Thus, apart from considerable improvement during last years, public services delivery remains the major challenge for local governments. Their websites will either include electronic distribution of services to online offer, or they stay behind the main current of information revolution.

VI. CONCLUSION

The use of ICT can support various forms of democracy. E-voting obviously strengthens representative, procedural model of government, while virtual local communities reinforce libertarian aspects of democracy [17]. In the theoretical framework of the research presented in this paper the ultimate model of democracy formed by ICT was not decided in advance. Instead, the general term of "citizens empowerment" was introduced, as the possible model of the growing role of individuals in political processes.

It is worth stressing that the same analytical framework can be applied in the study of various levels of government. Further research is needed however to assess if phenomena and processes which characterize local democracy are visible in states and democratic political systems as well.

The data presented above indicate that with regard to Polish local government websites we can barely observe the reinforcement of the individuals' status in their interactions with public institutions. Thanks to websites citizens have certainly much better access to information. Thus, in that domain costs they bear to pursue their democratic rights are visibly reduced. It is however the only dimension of e-government that really works. Well informed individuals, who are ready to get involved in local public life face the challenge of a very poor offer provided by local authorities. In case of Polish counties, the increase of e-participation is not among top priorities of local elites. Obviously, there are numerous locations where online local debate can proceed. Websites of local government (which are the natural spot to confront residents' and their representatives' views and opinions) are not in the forefront of civic engagement encouragement.

Electronic delivery of services, which makes operations of public institutions more transparent and customer oriented is yet another aspect of possible empowerment of citizens. This element of e-government has recently undergone perhaps the most spectacular transformation to enable the improvement of a citizen position in his or her relation with the state. Nevertheless, benefits of the ICT use in service distribution bypass users of Polish local government sites. Inability to perform transactions with the office and very limited access to online applications prove that electronic service delivery is still in the preliminary stage of development. In his or her relations with local administration a resident of the county is still considered more as the

 ISBN: 978-1-61208-176-2

passive petitioner than the customer, whose satisfaction is critical in the evaluation of the performance of the office.

The overall image of local e-government in Poland, based on data concerning counties websites, does not support the thesis of the observable reinforcement of the role individuals play in political processes. Thus far, ICT seem to have very limited impact on the nature of local democracy. They rather strengthen the existing rules of the game, with dominating position of political institutions, sluggish public debate and poor intensity of political participation. Members of local communities are still in search for the effective means to empower their political position. It seems that at the moment local government websites remain less than helpful in this endeavor.

ACKNOWLEDGMENT

This work was funded by the Polish Ministry of Science and Higher Education (research project N N116 331338).

REFERENCES

[1] D. Janssen, S. Rotthier and K. Snijkers, "If you measure it they will score: an assessment of the international eGovernment benchmarking," Information Polity, Vol. 9, 2004, pp. 121-130.

[2] F. Salem, "Benchmarking the e-government bulldozer: beyond measuring the tread marks," Measuring Business Excellence, Vol. 11, Issue 4, 2007, pp. 9-22.

[3] F. Bannister, "The curse of the benchmark: an assessment of the validity and value of a e-government comparisons," International Review of Administrative Sciences, Vol. 73, June 2007, pp. 171-188.

[4] L. Grossman, The Electronic Republic, New York: Viking, 1995.

[5] N. Negroponte, Being Digital, New York: Vintage, 1995.

[6] M. Margolis and D. Resnick, Politics as Usual: The Cyberspace 'Revolution', London: Sage, 2000.

[7] P. Norris, A Virtuous Circle: Political Communications in Postindustrial Societies, Cambridge: Cambridge University Press, 2000.

[8] L. Porebski, "Three faces of electronic democracy," Proc. Xth European Conference on Information systems (ECIS 2002), Gdansk, June 2002, pp. 1218-1227.

[9] K. Edmiston, "State and local e-government: prospects and challenges," American Review of Public Administration, Vol. 33, March 2003, pp. 20-45.

[10] L. Berntzen and M. Olsen, "Benchmarking E-Government: a comparative review of three international benchmarking studies," Proc. Third International Conference on Digital society, 2009, pp. 77-82.

[11] T. Carbo and J. Williams, "Models and metrics for evaluating local electronic government systems and services," Electronic Journal of E-Government, Vol. 2, 2004, pp. 95-104.

[12] A. Evangelidis, J. Akomode, A. Taleb-Bendiab and M. Taylor, "Risk assessment & success factors for e-government in a UK establishment," Proc. The First International Conference on Electronic Government, 2002, pp. 395-401.

[13] P. Ferber, F. Foltz and R. Pugliese, "The Politics of state legislature web sites: making e-government more participatory," Bulletin of Science, Technology & Society, Vol. 23, June 2003, pp. 157-167.

[14] P. Leith, and J. Morison, "Communication and dialogue: what government websites might tell us about citizenship and governance," International Review of Law, Computers and Technology, Vol. 18, Issue 1, 2004, pp. 25-35.

[15] J. H. Lim and S. Y. Tang, "Urban e-government initiatives and environmental decision performance in Korea," Journal of Public Administration Research and Theory, Vol. 18, January 2008, pp. 109-138.

[16] L. Porebski, "Evaluating the development of eGovernment dystems: the case of Polish local government Web Sites," Proc. the 11th European Conference on eGovernment (ECEG 2011), June 2011, pp. 475-481.

[17] J. Van Dijk, The Network Society: Social Aspects of New Media, London: Sage, 1999.

 ISBN: 978-1-61208-176-2

Community Detection based on Structural and Attribute Similarities

The Anh Dang, Emmanuel Viennet
L2TI - Institut Galilée - Université Paris-Nord
99, avenue Jean-Baptiste Clément - 93430 Villetaneuse - France
{*theanh.dang,emmanuel.viennet*}*@univ-paris13.fr*

***Abstract*—The study of social networks has gained much interest from the research community in recent years. One important challenge is to search for communities in social networks. A community is defined as a group of users such that they interact with each other more frequently than with those outside the group. Being able to identify the community structure can facilitate many tasks such as recommendation of friends, network analysis and visualization. In real-world networks, in addition to topological structure (i.e., links), content information is also available. Existing community detection methods are usually based on the structural features and do not take into account the attributes of nodes. In this paper, we propose two algorithms that use both structural and attribute information to extract communities. Our methods partition a graph with attributes into communities so that the nodes in the same community are densely connected as well as homogeneous. Experimental results demonstrate that our methods provide more meaningful communities than conventional methods that consider only relationship information.**

***Keywords*-social network; community detection; clustering;**

I. Introduction

Social networks of various kinds demonstrate a feature called community structure. Individuals in a network tend to form closely-knit groups. The groups are called communities or clusters in different context. Community detection is the task of detecting these cohesive groups in a social network [1] [2]. In many real-world networks, in addition to topological structure, content information is also available. Data is associated to the nodes and in the form of text, images, etc. For example in a social network, each user has information about age, profession, interests, etc. When content data is available, it might be relevant to extract groups of nodes that are not only connected in the social graph but also share similar attributes.

Many existing community detection techniques only focus on the topological structure of the graph. On the other hand, data clustering has been studied for a long time but most algorithms (e.g., k-means, EM) do not deal with relational data. The work of incorporating structural and attribute data has not been throughly studied yet in the context of large social graphs. This is the motivation of our work. Our key contributions are summarized next. In this paper, we study the relationship between semantic similarity of users and the topology of social networks (homophily concept). We propose two approaches to extract communities on several real-world datasets. Based on our evaluations, we conclude that our methods are able to discover more relevant communities.

II. Related Work

Detecting communities in a social network is still an open problem in social network analysis. In literature, many community detection methods have been proposed. According to [1], these approaches can be divided into four categories: node-centric, group-centric, network-centric and hierarchic-centric. Some popular methods are modularity maximization [3] [4], Givan-Newman algorithm [5], Louvain algorithm [6], clique percolation [7], link communities [8]. [2] and [9] provide a throughout review of the topic. However, these methods ignore the attributes of the nodes. Below are some studies that incorporate node attributes in the clustering process. Steinhaeuser et al. [10] proposed an edge weighting method NAS (Node Attribute Similarity) that takes into account node attributes. A community detection method is then proposed based on random walks. The complexity of the algorithm is $O(n^2 logn)$ (for random walks) or $O(n)$ (for scalable random walks) where n is the number of nodes. Zhou et al. [11] defined a unified distance measure to combine structural and attribute similarities. Attribute nodes and edges are added to the original graph to connect nodes which share attribute values. A neighborhood random walk model is used to measure the node closeness on the augmented graph. A clustering algorithm SA-Cluster is proposed, following the K-Medoids method. The time complexity of the algorithm is $O(n^3)$.

Coupling relationship and content information in social network for community discovery is an emerging research area because current methods do not focus on social graphs or they are not efficient for large-scale datasets.

III. Problem Statement

An attributed graph is denoted as $G = (V, E, X)$, where V is the set of nodes, E is set of edges, $X = X^1, ..., X^d$ is the set of d attributes associated with the nodes in V. Each vertex v_i is associated with an attribute vector $(x_i^1, ..., x_i^d)$. The goal of this work is to find communities in an attributed graph, that is to partition the graph into K disjoint groups (i.e., communities) $G_i = (V_i, E_i, X)$, where $V = \cup_{i=1}^{K} V_i$

 ISBN: 978-1-61208-176-2

and $V_i \cap V_j = \emptyset \; \forall i \neq j$. Nodes in the same communities are expected to be highly connected and have similar attributes.

Before clustering, a similarity measure must be determined. Our algorithms do not depend on the details of the measurement. Let $simA(i,j)$ be the similarity between a pair of nodes (i,j) in an attributed graph $G = (V, E, X)$. The measure should reflect the degree of closeness of the nodes in terms of their attribute values. An attribute can be classified as continuous, discrete or textual.

If the attributes are discrete, a commonly used similarity measure is based on the simple matching criterion. The similarity between two nodes in an attributed graph is determined by examining each of the d attributes and counting the number of attribute values they have in common.

For continuous attributes, the most commonly used metric is based on the Euclidean distance.

$$simA(i,j) = \frac{1}{1 + \sqrt{\sum_d (x_i^d - x_j^d)^2}}$$

If the attributes are textual, we first need to transform them into numeric values. A text document can be represented as bag of words. Each word is represented as a separate variable having numeric weight. The most popular weighting schema is tf-idf (term frequency-inverse document frequency). Each document is then represented as a vector of weight. To measure the similarity between two document vectors, cosine similarity is the most widely used metric.

IV. Community detection algorithms

In this section, we present two methods to discover communities in an attributed graph, given a similarity measure.

A. Algorithm SAC1

Our first approach is based on the modification of Newman's well-known modularity function. Given a graph of n nodes and m edges, $G_{i,j}$ represents the link (i,j), d_i is the degree of node i. If a graph is partitioned into K clusters, Newman's modularity [3] can be written as

$$Q_{Newman} = \sum_{l=1}^{K} \sum_{i \in C_l, j \in C_l} S(i,j) \qquad (1)$$

where the link strength $S(i,j)$ between two nodes i and j is measured by comparing the true network interaction G_{ij} with the expected number of connections $(d_i \cdot d_j)/2m$

$$S(i,j) = \frac{1}{2m} \cdot \left(G_{i,j} - \frac{d_i \cdot d_j}{2m} \right)$$

Newman's modularity does not include the attribute similarity between nodes. We define the "modularity attribute" Q_{Attr} of a partition as

$$Q_{Attr} = \sum_{C} \sum_{i,j \in C} simA(i,j) \qquad (2)$$

where $simA$ is the attribute similarity function.

Next, we introduce a composite modularity as a weighted combination of modularity structure (1) and modularity attribute (2)

$$Q = \sum_{C} \sum_{i,j \in C} (\alpha \cdot S(i,j) + (1-\alpha) \cdot simA(i,j)) \qquad (3)$$

α is the weighting factor, $0 \leq \alpha \leq 1$.

The next step is to find an approximate optimization of Q (direct optimization is a NP-hard problem [12]). We follow an approach directly inspired by the Louvain algorithm [6]. The algorithm starts with each node belonging to a separated community. A node is then chosen randomly. The algorithm tries to move this node from its current community. If a positive gain is found, the node is then placed to the community with the maximum gain. Otherwise, it stays in its original community. This step is applied repeatedly until no more improvement is achieved.

When moving node x to community C, the composite modularity gain is calculated as

$$\Delta Q = \alpha \cdot \Delta Q_{Newman} + (1-\alpha) \cdot \Delta Q_{Attr} \qquad (4)$$

in which

- Gain of modularity structure ΔQ_{Newman} :

$$\begin{aligned} \Delta Q_{Newman} &= \sum_{i,j \in C \cup x} S(i,j) - \sum_{i,j \in C} S(i,j) \\ &= \frac{1}{2m} \left(\sum_{i \in C} G_{i,x} - \frac{d_x}{2m} \sum_{i \in C} d_i \right) \end{aligned}$$

- Gain of modularity attribute ΔQ_{Attr} :

$$\begin{aligned} \Delta Q_{Attr} &= \sum_{i,j \in C \cup x} simA(i,j) - \sum_{i,j \in C} simA(i,j) \\ &= \sum_{i \in C} simA(x,i) \end{aligned}$$

The first phase is completed when there is no more positive gain by moving of nodes. Following Louvain, we can reapply this phase by grouping the nodes in the same communities to a new community-node. The weights between new nodes are given by the sum of the weight of the links between nodes in the corresponding communities [6]. To determine the attribute similarity between two communities, we propose two approaches. The first is to sum up the similarity of their members, the second way is to set to the similarity of their centroids.

B. Algorithm SAC2

Our first algorithm SAC1 repetitively checks all nodes, leading to $O(n^2)$ complexity. To reduce the computational cost, we propose another approach that only makes use of a node's nearest neighbors. Given an attributed graph:

 ISBN: 978-1-61208-176-2

Algorithm 1 Structure-Attribute Clustering Algorithm SAC1

Input: An attributed graph $G = (V, E, X)$ and a similarity matrix
Output: A set of communities
Phase 1 : Initialize each node to a separated community
repeat
for $i \in V$ **do**
for $j \in V$ **do**
Remove i from its community, place to j's community
Compute the composite modularity gain ΔQ
end for
Choose j with maximum positive gain (if exists) and move i to j's community
Otherwise i stays in its community
end for
until No further improvement in modularity
Phase 2
- Each community is considered as new node
- Reapply Phase 1

$G = (V, E, X)$, we define a k-nearest neighbor graph (k-NN) $G_k = (V, E_k)$ as a directed graph in which each node has exactly k edges, connecting to its k most similar neighbors in G. The similarity measure between 2 nodes i and j is defined as

$$S(i, j) = \alpha \cdot G_{i,j} + (1 - \alpha) \cdot simA(i, j)$$

where $simA(i, j)$ is the attribute similarity function, $G_{i,j}$ represents the link (i, j). Note that we can replace $G_{i,j}$ by other similarity measurements such as Jaccard similarity, cosine similarity, etc. [13] discussed several similarity metrics based on local information. Similar to the previous algorithm, we use α as a weighting factor.

We apply the measurement S in the first place to construct the nearest neighbor graph. In G_k, a structural edge represents the similarity between nodes (in terms of structure and attribute) in the original graph G.

The naive approach to build k-NN graph uses $O(n^2)$ time and $O(nk)$ space. However substantial effort has been devoted to speed up the process, such as parallel algorithms ([14], [15]), approximation algorithms ([16], [17]). In most recent work, [18] introduced $NN - Descent$, an algorithm for approximate k-NN construction with an arbitrary similarity measure. The method is scalable with the empirical cost $O(n^{1.14})$.

We propose a simple algorithm with two phases: constructing a k-NN graph G_k and finding structural communities in G_k to obtain the final clustering. In Phase 2, various methods can be employed to find communities. In our experiments, we choose Louvain as the detection method

Algorithm 2 Structure-Attribute Clustering Algorithm SAC2

Input: An attributed graph $G = (V, E, X)$
Output: A set of communities
Phase 1: Construct k-NN Graph G_k
Phase 2: Apply detection method to find structural communities in Gk. The result corresponds to the communities in G

because of its scalability. We set k equal to the average degree of the nodes in the graph G.

V. Experimental study

A. Experimental Datasets

We perform experiments to evaluate our algorithm on several real social networks:

Political Blogs Dataset: A directed network of hyperlinks between weblogs on US politics, recorded in 2005 by Adamic and Glance [19]. This dataset contains 1,490 weblogs with 19,090 hyperlinks between these webblogs. Each blog in the dataset has an attribute describing its political leaning as either *liberal* of *conservative*.

Facebook Friendship Datasets: The datasets contain the Facebook networks (from a date in Sept. 2005) from these colleges: Caltech, Princeton, Georgetown and UNC Chapel Hill [20]. The links represent the friendship on Facebook. Each user has the following attributes: ID, a student/faculty status flag, gender, major, second major/minor (if applicable), dormitory(house), year and high school.

DBLP Dataset: A co-authorship network with 10,000 authors, captured from the DBLP Bibliography data in four research areas: database (DB), data mining (DM), information retrieval (IR) and artificial intelligence (AI). Each author has two attributes: prolific and primary topic. Details of this dataset can be found in [11].

One of the most fundamental characteristic of social network is homophily [21]. The principle of homophily states that actors in a social network tend to be similar (i.e., to share some common attributes) with their connected neighbors, or "friends". In order to show this feature, for each attribute a in the dataset (e.g., political view, dormitory, year), we compute the probability that two friends are similar and compare to the probability of a random pairwise sample

$$P_{sl} = P(Similar|Link) = \frac{|(i, j) \in E : \ s.t. \ a_i = a_j|}{|E|}$$

$$P_s = P(Similar) = \frac{|(i, j) : \ s.t. \ a_i = a_j|}{|E| \cdot (|E| - 1)}$$

Table I shows that the similarities between friends are significant higher than random, according to a particular attribute. In Political Blogs, 90% of connected blogs are similar, compared to 49% of random pair. In Caltech network, similarity in dormitory are significant between friends (42%

Table I: Homophily measurement in experimental datasets

Graph	#Nodes	#Edges	Attribute	$\mathbf{P_{sl}}$	$\mathbf{P_s}$
Political Blogs	1,490	16,716	Leaning	0.90	0.49
Caltech	796	16,656	Dorm	0.42	0.12
Princeton	6,596	293,320	Year	0.53	0.13
Georgetown	9,414	425,638	Year	0.58	0.13
UNC	18,163	766,800	Year	0.43	0.15
DBLP	10,000	28,110	Topic	0.35	0.01

compared to 12%). In the graphs Princeton, Georgetown and UNC, friends are more likely to have the same class year. In DBLP, authors are most likely not connected if they do not share the primary topic.

The analysis of homophily demonstrates the correlation between structure and attribute information in real social networks. For that matter, node attributes could provide valuable information to facilitate community discovery.

B. Evaluation Measures

We extract the communities from the above datasets, using 6 different methods:

- Attribute-based clustering: K-means method is used to group nodes based on the similarity in attributes (link information is ignored).
- Random walks: Method proposed by Steinhaeuser et al. [10], based on random walks and hierarchical clustering. The walk length is set to the number of nodes.
- Louvain algorithm on unweighted graph.
- Fast greedy: Method proposed by Clauset et al. [22] based on the greedy optimization of modularity. The graph is weighted by node attribute similarities.
- Our proposed algorithms SAC1 and SAC2.

To evaluate the quality of these methods, we compare the number of communities, size of communities, modularity structure, modularity attribute and additional two measurements: density D and entropy E

$$D = \sum_{c=1}^{K} \frac{m_c}{m}$$

where m_c is number of edges in community c, m is the number of edges in G, K is the number of communities. D reflects the proportion of community intra-links over total number of links. High density denotes good separation of communities.

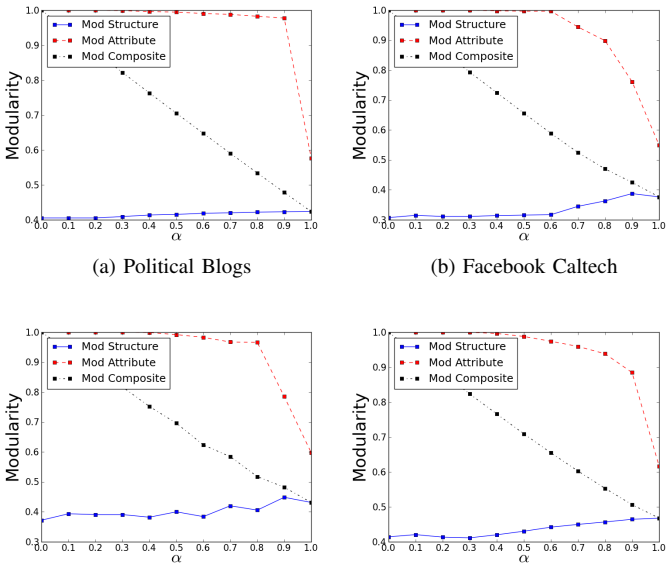
$$E = \sum_{c=1}^{K} \frac{n_c}{n} \cdot entropy(c)$$

$$entropy(c) = -\sum_{i} p_{ic}\ log(p_{ic})$$

where n_c is the number of nodes in community c, n is the number of nodes in G, p_{ic} is the percentage of nodes in c with attribute i. Communities with low entropy means they are more homogeneous with respect to the attribute a_i.

C. Comparison of SAC1 and SAC2

Because our approaches make use of the parameter α as a weighting factor between structural similarities and attribute similarities, we first examine the community qualities with different values of α. Figure 1 plots the modularity structure (E.q (1)), modularity attribute (E.q (2)) and modularity composite (E.q (3)) of SAC1's communities (in 4 graphs), for $\alpha \in [0, 1]$. The x-axis represents the values of α, the y-axis represents the modularities values. There is an increasing trend of modularity structure and decreasing trend of modularity attribute since the algorithm gives more favor to structural similarities as α increases. For SAC2 (not shown here), the modularities also follow the similar patterns.



(a) Political Blogs (b) Facebook Caltech

(c) Facebook Princeton (d) Facebook Georgetown

Figure 1: SAC1 modularity structure, modularity attribute and modularity composite for $\alpha \in [0, 1]$

Table II reports the average entropy and density of SAC1 and SAC2 on the datasets. Average entropy of SAC2 is lower than SAC1's whereas density of SAC1 is higher than SAC2's. That is, SAC2's communities are more homogeneous, but in terms of density, SAC1's communities are more dense.

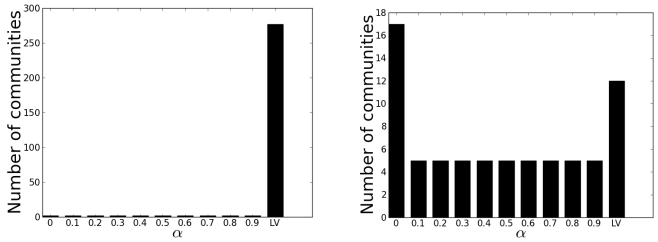
 ISBN: 978-1-61208-176-2

Table II: Average entropy and density of SAC1 and SAC2

Graph	Average Entropy		Average Density	
	SAC1	SAC2	SAC1	SAC2
Political Blogs	0.06	0.1	0.91	0.90
Caltech	0.75	0.33	0.50	0.46
Princeton	1.04	0.41	0.64	0.55
Georgetown	0.91	0.41	0.68	0.60
UNC	1.76	0.51	0.64	0.45
DBLP	3.01	1.24	0.82	0.52

D. Comparison against other methods

1) Number of communities and size distribution: We observe that SAC1 and SAC2 result in less number of communities than other methods. Figure 2 shows the number of communities found by Louvain and SAC1. The x-axis represent values of α, the y-axis is the number of corresponding communities. The outermost right bar is the number of communities from Louvain. It is clear that Louvain results in more communities. The result is similar for SAC2 (Table III). However, many of the communities found by Louvain are very small. For instance in Political Blogs, although 276 communities are found, the biggest two communities already consist of 80 percent of nodes. The rest of communities have the maximum size of 5 nodes. On the other hand, our algorithms correctly identified two communities in this graph, which correspond to two political views: liberal and conservative. It is observed that for large networks, Louvain often results in a few mega-sized communities and numerous small-sized communities. Our methods achieved a more balanced distribution of community sizes.



(a) Political Blogs (b) Facebook Georgetown

Figure 2: Number of communities in SAC1 (plot of α) and Louvain

Table III: Number of communities in SAC2($\alpha = 0.5$), Louvain and Fast greedy

Graph	SAC2	Louvain	Fast greedy
Political Blogs	2	277	277
Caltech	7	10	9
Princeton	7	20	24
Georgetown	9	12	42
UNC	7	19	31
DBLP	47	566	864

2) Community quality: Table IV and V compare the clustering entropy and density (with $\alpha = 0.5$) on two datasets. It shows that SAC1 and SAC2 result in communities with lower entropy (higher attribute similarities) than Louvain and Fast greedy's communities. For example, in Caltech graph, the entropy of SAC1 and SAC2 is 0.75 and 0.33 respectively, while the entropy of Louvain and Fast greedy is 1.65 and 1.71. On the other hand, the density of our methods is a little lower than the density of these two methods but higher than attribute-based clustering and random walks. For other datasets, the results are also similar.

Table IV: Entropy and Density of Caltech's communities

Method	Entropy	Density
Attribute-based	0	0.42
Random walks	0	0.35
Louvain	1.65	0.57
Fast greedy	1.71	0.56
SAC1	0.75	0.50
SAC2	0.33	0.46

Table V: Entropy and Density of Princeton's communities

Method	Entropy	Density
Attribute-based	0	0.53
Random walks	0	0.47
Louvain	1.71	0.62
Fast greedy	1.80	0.74
SAC1	0.84	0.62
SAC2	0.41	0.55

VI. Discussions

Both of our methods are parameterized, i.e., using α as a weighting factor, the natural question is how to choose α. Note that the results are quite stable with respect to α. With no domain knowledge, it is difficult to determine the value of α a priori. However, in social networks, we expect the links contain more information than attribute values. Based on this idea, we propose a strategy to approximate α. It is illustrated below:

init:

- $\alpha = 1$
- Set an interval i (e.g., $i = 0.1$ in our experiments)

repeat

- Compute the optimized clustering corresponding to α
- Let $Q_{Newman}(\alpha)$ and $Q_{Attr}(\alpha)$ be the modularity structure and modularity attribute of the partition
- Let $\alpha' = \alpha - i$
- Let $\Delta = (Q_{Newman}(\alpha') - Q_{Newman}(\alpha)) + (Q_{Attr}(\alpha') - Q_{Attr}(\alpha))$
- $\alpha = \alpha'$

until $\Delta <= 0$

Table VI reports the value of α found using the aforementioned strategy for SAC1 algorithm. It shows that the communities found are reasonably good in terms of modularity values.

Table VI: Optimum α found for the graphs

Graph	α	Q_{Newman}	Q_{Attr}
Political Blogs	0.5	0.41	0.99
Caltech	0.6	0.31	0.99
Princeton	0.7	0.42	0.96
Georgetown	0.5	0.43	0.98
UNC	0.6	0.33	0.91
DBLP	0.5	0.27	0.83

VII. Conclusion and perspectives

In this paper, we studied the issue of community detection in attributed graphs. We propose two methods that couple topological structure as well as attribute information in the detection process. Experimental results in real social networks demonstrated that our methods achieve a flexibility in combining structural and attribute similarities, hence able to bring in more meaningful communities. As future work, we try to bring further enhancements to our methods, e.g., reduce the algorithms' complexity, explore different similarity functions. We will apply our methods in different scenarios, for example with textual data or missing attribute values. We try to understand the roles of links and content information in the formation of online communities in order to devise adapted discovery strategies and to model the dynamic of the networks.

Acknowledgment

This work was partially supported by the projects ANR Ex DEUSS and DGCIS CEDRES.

References

[1] L. Tang and H. Liu, *Community Detection and Mining in Social Media (Synthesis Lectures on Data Mining and Knowledge Discovery)*. Morgan-Claypool, 2010, ch. 3.

[2] S. Fortunato, "Community detection in graphs," *Physics Reports 486, 75-174 (2010)*, 2009.

[3] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Physical Review E*, vol. 70, p. 066111, 2004.

[4] K. Wakita and T. Tsurumi, "Finding community structure in mega-scale social networks," *Computer Science - Computers and Society, Physics - Physics and Society*, 2007.

[5] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in network," *Phys. Rev. E 69, 026113*, 2004.

[6] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008 (12pp), 2008.

[7] G. Palla, I. Derenyi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," *Nature*, vol. 435, no. 7043, pp. 814–818, Jun. 2005.

[8] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann, "Link communities reveal multiscale complexity in networks," *Nature*, vol. 466, no. 7307, pp. 761–764, Jun. 2010.

[9] J. Leskovec, K. J. Lang, and M. W. Mahoney, "Empirical comparison of algorithms for network community detection," *CoRR*, vol. abs/1004.3539, 2010.

[10] K. Steinhaeuser and N. V. Chawla, "Identifying and evaluating community structure in complex networks," *Pattern Recognition Letters*, Nov. 2009.

[11] Y. Zhou, H. Cheng, and J. X. Yu, "Graph clustering based on structural/attribute similarities," *Proc. VLDB Endow.*, vol. 2, pp. 718–729, August 2009.

[12] U. Brandes, D. Delling, M. Gaertler, R. Goerke, M. Hoefer, Z. Nikoloski, and D. Wagner, "Maximizing modularity is hard," *ArXiv Physics eprints*, vol. physics.da, no. 001907, 2006.

[13] T. Zhou, L. Lu, and Y.-C. Zhang, "Predicting missing links via local information," *European Physical Journal B*, vol. 71, no. 4, pp. 623–630, 2009.

[14] M. Connor and P. Kumar, "Fast construction of k-nearest neighbor graphs for point clouds." *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 4, pp. 599–608, 2009.

[15] M. D. Lieberman, J. Sankaranarayanan, and H. Samet, "A fast similarity join algorithm using graphics processing units," *2008 IEEE 24th International Conference on Data Engineering*, vol. 25, no. April, pp. 1111–1120, 2008.

[16] J. Chen, H. Fang, and Y. Saad, "Fast approximate knn graph construction for high dimensional data via recursive lanczos bisection," *Journal of Machine Learning Research*, vol. 10, no. 2009, pp. 1989–2012, 2009.

[17] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching in fixed dimensions," in *ACM-SIAM SYMPOSIUM ON DISCRETE ALGORITHMS*, 1994, pp. 573–582.

[18] W. Dong, C. Moses, and K. Li, "Efficient k-nearest neighbor graph construction for generic similarity measures," in *Proceedings of the 20th international conference on World wide web*, ser. WWW '11, 2011.

[19] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 u.s. election: divided they blog," in *Proceedings of the 3rd international workshop on Link discovery*, ser. LinkKDD '05, 2005.

[20] A. L. Traud, E. D. Kelsic, P. J. Mucha, and M. A. Porter, "Comparing community structure to characteristics in online collegiate social networks," 2010, sIAM Review, in press (arXiv:0809.0960).

[21] P. F. Lazarsfeld and R. K. Merton, "Friendship as a social process: A substantive and methodological analysis," in *Freedom and Control in Modern Society*. Van Nostrand, 1954.

[22] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Physical Review E*, vol. 70, no. 6, pp. 1–6, 2004.

 ISBN: 978-1-61208-176-2

The Evolution of the e-ID card in Belgium: Data Privacy and Multi-Application Usage

Alea Fairchild
Vrije Univ. Brussel (VUB)
Brussels, Belgium
Alea.fairchild@vub.ac.be

Bruno de Vuyst
Vrije Univ. Brussel (VUB)
Brussels, Belgium
Bruno.de.vuyst@vub.ac.be

Abstract: **Since mandating in 2004 that all Belgian citizens carry electronic identification cards (e-ID), Belgium has been at the forefront of trends in electronic identification. As an e-ID card has become a necessity for service provisioning, the government has also started with distribution of e-ID cards to non-Belgians and children under the age of 12. Up until quite recently, the e-ID card only held the basic information of citizenship. This paper will examine the evolution of the e-ID card, and discuss the privacy issues of multi-application data on one card as the recent announcement of data for additional applications reopens the discussion of data linkage and data privacy for a card that is mandatory in usage.**

Keywords- e-ID; privacy; transparency; applications.

I. EVOLUTION OF BELGIAN E-ID

As this paper focuses on the evolution of e-ID in Belgium, we will not go into the concept of citizen acceptance of identity cards (paper or digital), as there are a number of papers that cover this from social and political aspects [1] [2] [3]. Acceptance has never been the issue, unlike in the UK [4], as Belgium has mandated the use of identity cards with the creation of a National Register of natural persons in 1983 [5]. The register issues unique identifiers for each Belgian citizen in the form YYMMDDNNNCC, where YYMMDD refers to the date of birth of the citizen, and NNN is an even number for females and odd for males. CC is a checksum so errors can be detected when processing the number automatically. The register also keeps track of current and past addresses and keeps a record of all the citizen's identity-related documents: passport, driving license and other relevant data. So citizens and residents cannot opt-out, but must carry the e-ID for identification and for service to be provided [4]. So choice is not part of our discussion.

In terms of acceptance, Belgians have already been used to showing national ID cards for identification for services, but the use of these paper-based ID cards have been on an event-oriented basis, and these respective events have not been tracked on a longitudinal basis. For example, a citizen may make a photocopy of his ID card for his bank to open a bank account. But that particular event of opening the bank account using the ID card is not recorded on a digital format in a public data facility where someone can use this event on a longitudinal basis [6].

For reasons of both efficiency and service provisioning, Belgium decided in early 2000 to be an early adopter and to trial the concept of a digital version of the paper-based national ID card. In Spring 2003, a pilot project across 11 municipalities was a trial of the e-ID and its implementation. In Spring 2004, the Belgian government decided, after approval of its legislative body, to mandate the e-ID for the whole country. A timeline of the e-ID implementation can be seen in Figure 1.

Although at the e-ID launch, there was only one application one can use the card for, ôTax-on-Webô [7], it was envisioned that this will be the basis for future service provisioning for several layers (federal, regional, local) of Belgian governance. However, this was initially concerning to Belgians because of ódigital trailsô: how the numerous events of usage of their ID card are used, and by whom.

This paper will examine the transparency of what data is held on the card, as who has the right to use/view that data have been concerns for the changeover from paper-based to digital ID cards. We discuss what data is held on the e-ID card, and what applications are available for e-ID at present in Belgium. We end with an explanation of privacy and transparency in Belgian e-ID cards, and how multi-application usage may be in the near future of this card.

II. E-ID IDENTITY DATA

The initial paper-based document included the following pieces of information on the citizen: name (family name, up to two given names, and the initial of a third name), address, title, nationality, place and date of birth, gender, and a photo of its holder.

For the e-ID card, it is visually similar to the previous identity card and shows the same information as the paper-based document, except for the address. It also contains a hand written signature of its holder and also of the civil servant who issued the card. It also mentions the validity dates of the card (the card is valid for five years), the card number, the national number of its holder, and the place of delivery of the card [8].

All this information is also stored on the chip in a so-called ñidentity fileò. The identity file is around 200 bytes long, and is signed by the National Register (RRN). In addition to the identity file, there is also an address file (about 150 bytes). This address file is kept independently as

 ISBN: 978-1-61208-176-2

the address of its holder may change within the validity period of the card. The RRN signs the address file together with the identity file to guarantee the link between these two files. The corresponding signature is stored as the address file's signature. As biometric feature, Belgium decided to use a photo (3 KBytes, JPEG format). This photo is (indirectly) signed through the RRN, as its hash is part of the user's identity file [8]. An example of a Belgian citizen can be seen in Figure 2. Cards for kids and for foreigners look differently.

Kids cards contain a unique safety feature to contact parents in case of emergency. This feature allows third parties to enter a list of preset phone numbers by way of a unique phone number and the childôs RRN, both visible on the kids-ID card. If a child is injured the parents can then be easily be contacted. The childôs parents are the ones that determine the preset list of phone numbers via a secure online database [9].

For usage of the e-ID card, there have been initial teething pains, with police cars needing to be equipped with readers in their glove compartment so people stopped for a possible violation can have their e-ID card read. When first issued, citizens had to carry an extra piece of paper with the e-ID card, as a police officer without a reader could not see the address of the citizen, which is one of the fundamental pieces of information requested from the ID card [6].

III. DATA HANDLING AND E-ID BENEFITS

The basic identity data are now digitally included in a microchip on the identity card, with a reader mechanism that allows a person to identify himself digitally and to place an electronic signature using the card and a password. In this way, storage and usage of citizensô data becomes a bit more user-centric.

According to the National Register [10], the benefits of an e-ID card are:

- Self identify on Internet;
- 24/7 availability of particular documents via Internet;
- Ability for the card holder the possibility to check information regarding themselves, that in the register or in the Rijksregister of the natural persons stands, to consult, and in order to know, which authorities, institutions and persons during the last six months have consulted or improved have, with exception of the municipal and judicial authorities, that entrusted are with the investigation and the repression of punishable facts;
- A protected electronic connection, online information exchange with the authorities or with private enterprises;
- A protected manner via the Internet commercial operations export, as well as a buyer as in the quality of seller (online buy and sell);
- Via the web numerous forms fill in: load declaration, request of a study appropriation or of an excerpt from the register;
- Through self to identify, get entry to various places: container park, building of an enterprise, library, sporthal é .;
- Mails sign or recorded send mails.

The development team at FEDICT (the Federal ICT office) has developed add-ons in Mozilla and other browsers to enable the citizen use of e-ID. It is already supported on several operating systems, including Linux (Open SUSE).

IV. MULIT-APPLICATION E-ID CARDS

Having an application that is mission critical to the citizen is one driver to get users to want to switch from paper to digital form. By promoting government applications such as ñtax on webò, registered mail, social security registration of new personnel, online consultation of government data, as well as the distribution to twelve-year olds of a free smart card reader when they get their e-ID card, the home penetration with readers was expected to increase in the short term [9].

These new applications offered by use of the eID are expanding, and the government feels that it will create a surplus value form for the citizen and for the concerned authority. But it is not clear if the citizens feel the same way, especially with data linked for different applications on the same card.

In August 2011, it was reported in the press [11] that the amount of personal information being stored on the compulsory Belgian ID card is being extended. In future personal social security information will also be stored on the e-ID card. The SIS (social security) card is being discontinued. Within the next year, pharmacies and doctors will start using e-ID cards instead of the SIS card in order to obtain personal social security information about their patients and customers. The two systems will operate in parallel for a while, and then the current SIS card will disappear by the end of 2013. The social security authorities and health insurance bodies are already paving the way for the switch-over [11]. At the time of the e-ID initial launch, although it was technically feasible to integrate the SIS data, at the political level it was considered to be too high an infringement on personal privacy and the integration was blocked [9]. The question might be at this time, what has changed? The willingness of the public, or the need for cost efficiency? As the public cannot opt-out, the sensitivity of the use of e-ID for multiple data applications needs to be of concern to the government.

 ISBN: 978-1-61208-176-2

V. PRIVACY AND TRANSPARENCY

Unlike the Austrian e-ID, which from the onset has attempted to be privacy-friendly through the use of unlinkability schemes, the Belgian e-ID card has not addressed any aspects of privacy such as unlinkability, or anonymity, as discussed by Pfitzmann et al. [13]. At present no other national e-ID card design scheme in Europe puts emphasis on privacy beyond data protection and retention [4].

Where event-based transaction data is retained in identified form, it can result in a collection of data that reveals a great deal about the individual and their behaviour. Such 'data trails' may be used to trace back over a person's past, or analysed to provide an abstract model of the person, or 'digital persona'. This digital persona may then used by government agencies as a means of social control, for example.

Since the 1980s, basic mechanisms for privacy-enhancing identity management under control of the user have been proposed [12] [13] [14]. Control by the user requires that he firstly knows about actual and potential processing of his personal data and secondly that he in principle can decide case-by-case on data disclosure to specific parties, possibly in the limits given by law and society. The most effective, yet not always realistic way to protect one's privacy is data minimisation, i.e., to disclose as little personal data as possible.

From a privacy point of view, the main issue in addition to the data-minimisation principle is the purpose-binding principle ï data should only be collected and used for a specific purpose.

In Belgium, the citizen can ask for and use his data, see if they are correct and see who has used them, as government workers also have to use their own e-ID card to provide the service. There is a website maintained by FEDICT, the national IT organization, that allows citizens to track their national ID number and who has been using it for what purpose. With their e-ID card, the citizen can open the data cabinet in which his data are safely stored with his identification key. The citizen can verify the data, eventually ask for correction, use his data, and see who, and at what time, entered the data cabinet. This level of transparency of the process and privacy authentication has been important in the enforced uptake.

A data privacy error was made in Belgium by including the structured register number in the certificates stored in the electronic ID card. This is something that must be avoided: the number leaks too much personal information about the citizen; in this case, age as the register number uses the date of birth in the number.

The only biometric included on Belgian e-ID cards is the holder's photo, which is about three kilobyte in size and not suited for automatic recognition of the cardholder. Correct implementation of biometric features is a very complicated issue, and may not be realistic and cost-effective. The Belgian eID card costs about ú12.50, including the chip, maintenance of the infrastructure and two certificates per cardholder with a validity of five years [8].

The authorities have switched over on 17 October 2008 to the production of e-ID cards on the New Belgium Root certificate. As each e-ID card has been initialized with a genuine copy of the Belgian Root CA certificate, the e-ID card can be used as a ñtrusted sourceò as users can verify the chain of trust within the Belgian PKI system by loading the Belgium Root CA certificate from her/his smart card. Apart from revoking the use of an e-ID card's keys when it is stolen, card holders also have the possibility to have the electronic signature capability of an e-ID card revoked, even before using a card [8].

VI. CONCLUSIONS ON THE FUTURE OF MULTI-APPLICATION DATA

The data from the SIS card will add information on the kind of health insurance the citizen holds (and that the citizen is insured, which is required in Belgium). Health records are not stored on the e-ID card. But the e-ID card is the linkage to the Crossroad Banks of Belgium, which are internal governmental information brokers on social security status, business information and car registration.

The question remains how cross linking of data may be used in a manner not fit for purpose, and what kind of legislation or audit trails will be utilized to protect citizen data privacy going forward as the government pushes to add multiple application usage to the e-ID card.

By promoting government applications, registered mail, social security registration of new personnel, online consultation of government data, together with the distribution to twelve-year olds of a free smart card reader when they get their e-ID card, the home penetration with readers is expected to increase in the short term.

However, privacy in a technological sense has not yet been included in the current version of the e-ID card. Belgian reliance on e-ID as a form of authentication and access means that no one can opt-out of the scheme, which makes security, transparency and privacy paramount to longer term interoperability within the EU. As new applications are added to the card, Belgians may get more wary of what can and cannot be linked together on the same card.

REFERENCES

[1] Walker, G. de Q. ñInformation as Power: Constitutional Implications of the Identity Numbering and ID Card Proposalò, *CIS Policy Report*, (1986) Vol. 2 No. 1, Centre for Independent Studies, St Leonards, NSW, February.

[2] Clarke, Roger ñHuman Identification in Information Systems Management Challenges and Public Policy Issuesò, *Information Technology & People*, 1994, Vol. 7(4), p. 6-37.

[3] UK Performance and Innovation Unit (PIU), óPrivacy and Data-sharing: The way forward for public servicesô, published by the PIU in April 2002.

[4] Arora, S. National e-ID card schemes: A European overview, Information Security Technical Report (2008) Vol. 13, p. 46-53.
[5] Belgisch Staatsblad , 28 March 2003. Retrieved 3 Dec 2011 at: http://eid.belgium.be/nl/binaries/WT1_tcm147-9980.pdf
[6] Fairchild, A and de Vuyst, B. Privacy, Transparency and Identity: The Implementation of the e-ID card in Belgium, *Proceedings of the 9th European Conference on Electronic Government* (ECEG 2009), University of Westminster, London, UK, June 29-30.
[7] Tax-on-Web, Federal portal, Retrieved 3 Dec 2011 at: http://www.tax-on-web.be
[8] De Cock, Danny; Wolf, Christopher; and Preneel, Bart (2007) The Belgian Electronic Identity Card (Overview), Working Paper of ESAT/COSIC, KU Leuven.
[9] Marien and Van Audenhove. The Belgian e-ID and its complex path to implementation and innovational change, *Identity in the Information Society,2010,* Vol. 3(1), p. 27-41.
[10] Rijksregister, Retrieved 26 Nov 2011 from URL: http://www.ibz.rrn.fgov.be/index.php?id=141&L=1
[11] FlandersNews.be. More personal info on electronic ID card soon. (20 August 2011). *deredactie.be*. Retrieved September 3, 2011, from http://www.deredactie.be/cm/vrtnieuws.english/news/1.1091792
[12] Chaum D. Security without identification: transaction systems to make big brother obsolete. Communications of the ACM, vol.28; Oct. 1985. no. 10, p. 1030ï 1044.
[13] Pfitzmann B,Waidner M, Pfitzmann A. Secure and anonymous electronic commerce: providing legal certainty in open digital systems without compromising anonymity, IBM research report RZ 3232 (#93278) 05/22/00 computer science/mathematics. Zurich: IBM Research Division.
[14] Leenes R, Schallabock J, Hansen M, editors. PRIME white paper, Retrieved 3 Dec 2011 from: https://www.prime-project.eu/prime_products/whitepaper/; 2007.

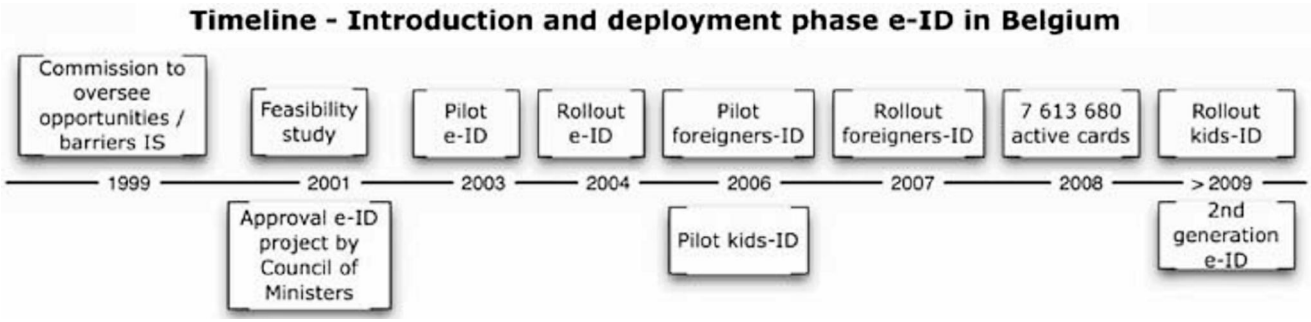


Figure 1. Timelineð Introduction and development of eID in Belgium [9]

Figure 2. Belgian eID card's visual aspects [8]

Analyzing Social Roles using Enriched Social Network on On-Line Sub-Communities.

Mathilde Forestier, Julien Velcin, Djamel A. Zighed
Eric Laboratory, University of Lyon
Lyon, France
mathilde.forestier@univ-lyon2.fr,
julien.velcin@univ-lyon2.fr,
abdelkader.zighed@univ-lyon2.fr

***Abstract*—Analyzing the social roles inside on-line communities became a big challenge nowadays. The on-line communities formed around exchange platforms (e.g., forums) create an increasing source of data for analyzing user's behavior. This paper proposes an exploratory analysis of communities in news website based on its sub-communities. Actually, we assume that people who participate in forum debate in news websites focus their participation in one or a very few topics (also called context), i.e., they formed the sub-communities. These sub-communities, will help us to find the *contextual celebrity*: the pertinent users in the sub-communities. We based our analysis on a dataset composed by 11,143 users writing more than 35,000 posts on 57 different forums grouped in 3 topics, and on social networks enriched with relations extracted from the content of the users' posts.**

***Keywords*-Social role; Social network; On-line community.**

I. INTRODUCTION

During the Roman era, the forum was the public place of the city, i.e., the social, political and economic center. The forum allowed people to communicate, exchange, debate and socialize. Forums still exist nowadays in a different way: thanks to the forums on the Web 2.0, users communicate interactively on a common interest.

People who participate in these forums (also called users) form an *on-line community*. Schoberth et al. [1] use this term "to describe the communication and social interaction that is seen in the Internet and web-based list servers, bulletin boards, Usenet newsgroups and chats". We can complete the definition with the one given by Hymnes [2] about the speech community which represents "a group of people who share rules for the conduct and the interpretation of speech, and rules for the interpretation of at least one linguistic variety". So, people who participate in forums form an on-line speech community. People in this on-line speech community, as in the real life [3], play a social role, as define in [4] "beside having personalities, by being part of the social group, people occupy positions in the social structures of groups that allow them to do and say certain things, as well as constrain them from saying or doing other thing". Golder and Donath follow the Goffman's theory [3] through which a role represents the "rights and duties attached to a given status". Still, in a Goffman's position, Gleave et al. [5] specify that a social role can only be apprehended in the interaction, i.e., people play a role depending on others.

In this paper, we focus on the understanding of sub-communities (from a whole community) in order to find good clues to comprehend the *contextual celebrity* social role. We define a sub-community as a sub-part of a whole community depending on a topic (also called context). In other words, a sub-community represents all the users who participate in a specific topic in a news website (e.g., politic, media, etc.). We assume that users participate in one a very few topics in their interest. So, the *contextual celebrity* represents a user particularly interested in a specific kind of topic compared to the whole on-line speech community. This user is recognized as a pertinent one by the other members of his sub-community.

So, the contributions of this paper are to explore an on-line community based on the analysis of the sub-communities which belong to it. The general idea is to confirm that users participate depending on a context and find some clues to detect the *contextual celebrity* for each kind of sub-communities. Note that, in this paper, we use the term topic or context independently.

This paper is organized as follow: first, we explain some related work and we position our work according to the existent one. Then, we describe the dataset we use to make the analysis of the social role inside sub-communities. We continue by briefly presenting the construction of our enriched social network using the structure and the content of the data. Finally, we explore the on-line speech community with its sub-communities and the concept of *contextual celebrity* social role.

A. Related Work

The social role analysis was highlighted by Goffman in [3]. According to his theory, human being adopts a "pre-established pattern of action, which is unfolded during a performance and which may be presented or played through on other occasions". According to him, individuals play a role during the interaction. This notion had a great

 ISBN: 978-1-61208-176-2

repercussion through the apparition of Web 2.0 and the emergence of new media of exchange. Some researchers used database of email exchange and probabilistic model as blockmodel to define some social roles in firms [6][7][8]. Other researchers looked at predefined roles as the expert [9] (who is the most expert?) or the influencers in social network [10][11] (who gets the power to convince people in the social network). In an other perspective, computer scientists and sociologists found a great interest to analyze the social roles in forum debates. Their works aim to extract social roles as a predefined behavior in the on-line speech community using a social network analysis and the user participation behavior. This double analysis allows to capture the place of the user inside the community based on his implication and his reputation. Golder and Donath made an ethnological study and found out six kinds of social roles: the celebrity, the newbie, the lurker, the flamer, the troll and the ranter (refers to [4] for the definitions). These social roles can be positive (e.g., the celebrity) or negative (e.g., the flamer or the troll). This ethnological approach considers that a content analysis of the posts brings a lot of informations. In our work, we use a content approach to extract our social network with the aim to define the social roles. We will see in Section I-C how we enrich our social network with new relations extracted from the content of the discussion. Others social roles have been discovered in this on-line speech community as the answer people and the discussion people [12]. In a political discussion context, Himelboim et al. [13] looked for the discussion catalyst. This kind of users influences the information that enters in a newsgroup and affect the discussion evolution within it. Kelly et al. [14] found three social roles in this kind of discussion: the friends, the foes and the fringes. The authors highlighted that people prefer to speak to users who are in another political affiliation than themselves. The great majority of the users in political discussion looks for virulent debate on society and way of life. Furthermore, the authors found the fringe social role which refers to a marginal group of people that raises interesting questions for qualitative study. Fisher et al. [15] took more largely into account the context of participation to analyze social roles. According to them, the user's participation is different if he participates in help forums opposed to a flame forums. Their idea makes us think that in Usenet, there are some specialized forums for flame, for help etc. But in a news website the configuration of participation is quite different, there is no specialized forum as in Usenet, but there are some topics where users are more interested to debate in. Very close to our work, Angeletou et al. [16] and Chan et al. [17] explain some on-line sub-communities by their composition of users roles, but each sub-community represents one community: there is no overlapping, no confrontation between the sub-communities. In this paper, the context of participation is represented by the topic which the forum belongs to, e.g., politic, media, living, etc. So, we propose a new way to understand social role depending on the context in on-line sub-communities. Finally, we refer the reader to Gleave et al. [5] and Forestier et al. [18] in order to have a larger state of the art and analysis about social roles.

B. Dataset introspection

In this section, we present the data we used to analyze the sub-communities and the *contextual celebrity*. We based our analysis on the forums of the HuffingtonPost (www.huffingtonpost.com) news website. We extracted 57 forums dealing with three topics, i.e., context: politic, living and media. The dataset is composed of 19 forums of each topic. The whole dataset contains 11,443 users and 35,175 posts. Table I presents the basic statistics on each topic. The overlapping of the sub-communities implies that the sum of the users from the three sub-communities is upper than 11,443 users. Note that the on-line speech community represents all the users and we are looking to the *contextual celebrity* in sub-communities (communities depending on a context).

Table I
BASIC STATISTICS ABOUT THE PARTICIPATION ON THE THREE TOPICS

	Politic	Living	Media
# of users	4547	3667	5973
# of posts	12725	8274	14176
Average number of posts per user	2.8	2.3	2.4
% of users who exclusively participate in this kind of topic	58%	68%	65.5%
% of users having one post on all users who participate in this topic	50%	58%	54%
% of users having between]1,5] posts	39%	34%	38%
% of users having between [6,11]	7%	5.7%	5.5%
% of users having between [12,16]	2%	<2%	<2%
% of users having between [17,∞[	<2%	<1%	1%
Total	100%	100%	100%

Table I shows that it exists in all sub-communities a hard core of specific users, i.e., users who participate only in one topic. Furthermore, the ratio between the number of posts and the number of users is quite the same in the three sub-communities. Users in living topic (respectively media topic) post an average of 2.3 messages (respectively 2.4). In politic forum, the ratio is a little bigger, i.e. 2.8, posts per user. Most of the users concentrates their participation on one topic and for each sub-community, at least half of the users post just one post in one topic. This comportment seems really interesting and, even if this is not the object of this paper, and in a perspective way, the study the behavior of these users through the sub-communities and in a temporal way, can be really interesting.

 ISBN: 978-1-61208-176-2

The three sub-communities follow the same rule of participation: most of the users posts less than six messages. There is a real gap between people who write less than six messages and those who write more. For each sub-community, an average of 6% of the user post between five and ten messages. Finally, a very few users posts more than ten messages in a topic. The *contextual celebrity* is being more likely among them.

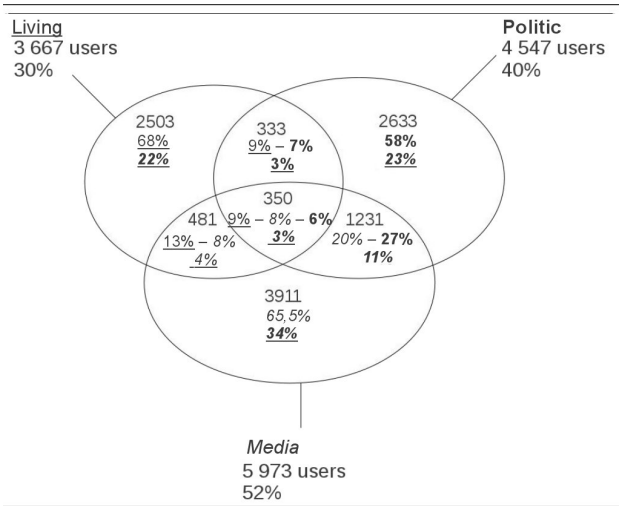


Figure 1. Overlapping of the tree topics

Figure 1 shows the overlapping of the sub-communities. The bold numbers represent the attributes of the politic sub-community, the underlined numbers represents the attributes of the living sub-communities and the ones in italic represent the attributes of the media sub-community. So, the numbers underlined, in bold and in italic represent the percentages for the whole community. The living (underlined on Figure 1) sub-community represents 30% of the community. Inside the living sub-community, 68% of the users participates only in this topic (it corresponds to 22% of the whole community). The statistics are quite the same in the two others sub-communities. A very few users participates in living topic and an other topic (9% of the users in living sub-community participates also in the politic topic (in bold), and 13% in both living and media topics (in italic). This minority represents only 3% and 4% of the whole community. Note that less than 30% of the media sub-community wrote at least one post in politic. These two sub-communities seem to be closer than the politic and living sub-communities. Finally, only 3% of the population participates in the three topics (it represents 9% of the users of living, 8% of politic and 6% of media sub-community). In conclusion, only 21% of the users participate in more than one topic.

C. Enriched social network

Web forums have the particularity that they structure the debate. Users who participate can, using this structure, reply to the post they want to reply to. This structure is used to extract social networks in existent works treating social roles [9][12][13][15]. But, reading the forums shows that people not only interact using the structure (reply to) but also through quotations. We find two kinds of quotations: the text quotation and the name quotation [19]. These two quotations allow an user to reply to several ones through one post; and people who read the forum automatically understand when an author is quoted (by the name, or by the quotation of a previous post). The idea is when a person quotes the name of another one, he adopts certain community codes and he considers himself to be entitled to refer to the person by his pseudonym, e.g., a newbie (i.e., new user) never feels the right to call other users by their pseudonym. So quoting the name implies the user's integration in the online speech community. The text quotation relation brings some important information during the analysis. Actually, more a user quotes another, more these two users are linked. Furthermore, the text quotation frequently implies a precise conversation between the two users, i.e., if I quote a part of your post, I really reply to you, and in most of case I argue your discourse with an opinion. To make a finer interaction analysis, we wanted that the analyze taking into account these quotations in a an automatically way. So, we created an enriched social network where users can be linked by three relations:

- The structural: a user replies to another one using the structure of the forum;
- The name quotation: when a user quotes the name of an other user in his post;
- The text quotation: when a user quotes a part of a previous post in his post.

Finally, we have three separates but complementary social networks, i.e., one social network for one kind of relation. Each of this social networks give some clues to understand the user behavior. The social network constructed with the name quotation relation gives some informations about the user's reputation: is he known by his sub-community? Is he often quoted? Is he often quotes? The social network constructed with the text quotation relation give some others clues: does the user like to debate? Does he bring some interesting informations to debate?

Our model reaches a quite good score in term of precision (ratio between number of quotations found by both evaluators and system compared to the number of quotations found by the system). We refer the reader to [19] for more information about the social network extraction.

Figure 2 shows the three separate social networks. We used the Jung Java toolkit for SNA to build the graphs. The social networks on Figure 2 are built only with users having written more than 15 posts in all the dataset. The gray scale of nodes represents the kind of forum a user participates in. Black nodes make the connection between subgraph of gray

 ISBN: 978-1-61208-176-2

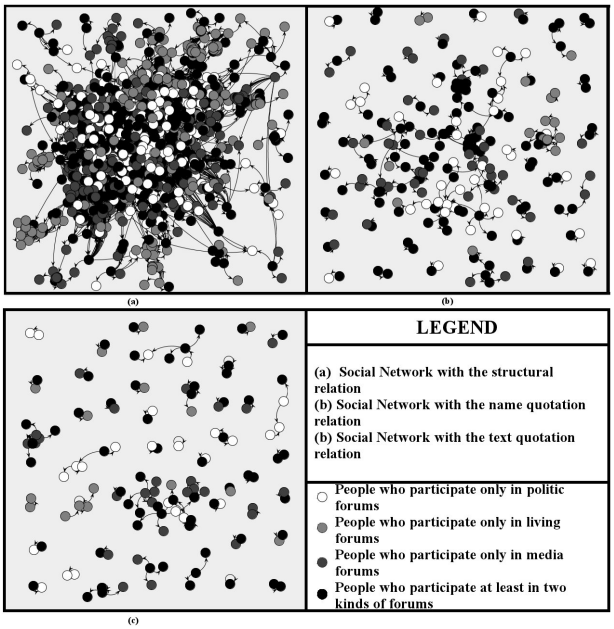


Figure 2. Social networks visualization

users (users who participate only in one kind of forum). We can also see that name quotation (b) is used more than the text quotation relation (c) by user who participate in the whole website. Referring to Table II, the name quotation represents the double of the text quotation (994 name quotations against 476 text quotations). The media sub-community uses more the name quotations compared to the two others sub-communities, we can interpret this result as a closer sub-community, i.e., the users of the media sub-community know better the others who participate in it. Surprisingly, it is not the users who post the most in a topic who quote or are quoted the most. This result is important concerning the detection of the *contextual celebrity*: users who are quoted by the name and the text and who post a lot of messages have more chance to be recognized by the others and to have a good reputation [20] in their sub-communities.

II. A NEW SYSTEM TO ANALYZE SUB-COMMUNITIES AND SOCIAL ROLES

As we saw in Section I-A, social role analysis became an important research study. Nowadays, it seems really important to understand who is who in the on-line speech communities. But, as we saw before in these works, most of the researchers use Usenet to extract social roles. The fact is that forums on news websites become increasingly generic while Usenet is quite specific. Furthermore, news websites allow users to treat several kinds of forums, e.g., politic, societal, etc. and the social role is dependent of the context[3][5]. The goal, here, is to retrieve social roles depending on the context, i.e., the kind of forum treated. Finally, the three relations between users (see Section I-C) allow a finer perception of the interaction. These relations will help us to a better extraction of the social roles.

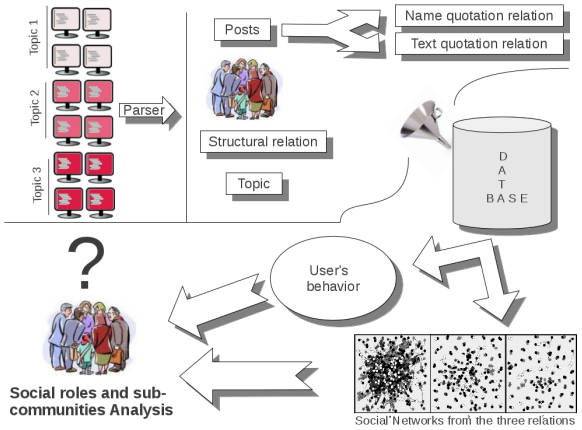


Figure 3. Presentation of the system

Figure 3 shows the process of our system from the website to the analysis. First of all, we collect the forums from the website using a parser. Note that the parser is specific for each website. The parser retrieves the forum topic, the users pseudonyms, the posts and the structural relation, i.e., which post replies to which one? Who replies to whom? Then, the system analyzes the content of the posts in order to extract the name and the text quotation relations. All the data is scored in a database. Finally, using the enriched social network (with the relations extracted from the content of posts) and the user's participation behavior, we analyze the social role based on the context, i.e., the topic of the forum.

We will present in the next section the way we choose to analyze the social roles taking into account the context and

Table II
BASIC STATISTICS ABOUT THE ENRICHED SOCIAL NETWORK ON THE THREE TOPICS. TQ : TEXT QUOTATION, NQ : NAME QUOTATION

	Politic	Living	Media
# of TQ	177	146	153
# of users who use TQ	151	119	118
# of NQ	350	183	461
# of users who use NQ	256	128	118
# of users having used TQ	151	119	118
# of users having more than 15 posts and are quoted by text	17	15	13
# of users having more than 15 posts and quoting by text	19	0	19
# of users having used NQ	256	128	375
# of users having more than 15 posts and are quoted by their name	33	15	23
# of users having more than 15 posts and quoting by the name	28	7	28

the kind of relations between users.

III. Sub-communities and social role Analysis

To analyze the sub-communities and to find the *contextual celebrity(ies)*, we decided to perform a principal component analysis (PCA)[21]. This unsupervised method aims to create a new description space of the data. We use Tanagra [22] to compute the PCA.

A. Criteria of analysis

We created several criteria to analyze the on-line community in a contextual perspective. These criteria are based on the individual's behavior and the analysis of the social network. We calculate for each individual who participate in the forums:

- Number of politic / living / media forums the user participates in;
- Number of posts in politic / living / media forums;
- In-degree with the structural relation function of the topic;
- Out-degree with the structural relation function of the topic.

So, each user is defined by 12 criteria measuring the user's interest in the topics and his place inside the sub-communities. Actually, the participation is comprehended by the user's participation behavior metrics; and his place inside the sub-community by his place in the social network using the in-degree and out-degree with the structural relation. These criteria allow us to create an unsupervised method to explore the on-line speech community.

B. Principal Component Analysis

The Principal Component Analysis (PCA) consists in transforming the criteria of analysis (see Section III-A) in new variables, each independent of each other. The aim of this method is to create a new space where the dimensions are not correlated one to the other. It also allows to reduce the information description to a limited number of components, less than the initial number of criteria of analysis. PCA is really interesting for several reasons. First of all, we want to explore the sub-communities in a unsupervised way. The social role of the users depends on the interest of the user for one topic and his place inside the sub-community. We expect that PCA finds three groups (one group for each topic) that are not correlated one to the others. Furthermore, this is an unsupervised method of analysis because our dataset does not allow the usage of supervised methods: we do not have labels to learn rather we have to discover and interpret the knowledge from the data. Finally, this old method (more than one century) made proof of its performance and it is still used today.

The first three axes found by the PCA resume 75% of the knowledge contained in the data. The fourth axis only adds 5% of supplementary information, so we keep the first three axes. Note that a resume of 75% of information is a quite good score for real data.

The first axis is described on the positive part by the politic topic: number of posts, in- and out-degree with the structural relation. On the negative side, the axis is described by the living topic. The second axis is constructed on the positive part with the politic forum. The third axis is constructed with all the criteria concerning the media topic. This construction proves that the on-line speech community is divided into sub-communities function of the topic. Nevertheless, the sub-communities are not completely separate (otherwise the PCA gives some correlations about one) and some users being part of several sub-communities.

Figure 4 represents the correlation scatter plot created with the two first axis of the PCA. The three forums are visibly separated. We have on the top left of the graphic the forums about living, on the bottom right the forums dealing with media and on the top right the forums dealing with the politic. This graphic proves that individuals have certain behavior depending on the kind of forum they participate in. Figure 4 shows that the angle between the living topic and the media topic is about 180° function of the gravity center (see the right line on figure 4 between the two groups). It means that it exists a negative correlation between the two groups. In other words, the more the users participate in forums dealing with media topic, the less they participate in living topic and vice versa. In another way, the politic topic is almost on a right angle compared to both media and living forums. There is a statistical independence between the politic topic compared to the media and the living topic. The PCA does not find a correlation between them. It seems that the participation in the politic topic does not influence the participation in media and/or living topic.

Finally, the PCA confirms that users mostly participate in one kind of topic (i.e., in a context). To find who are the *contextual celebrities* we propose to find users who maximize all the criteria on one topic and who have no or very few participation in the others. So, in a perspective way, we are thinking to use multicriteria aggregation so that we find not just one *contextual celebrity* per topic but a list

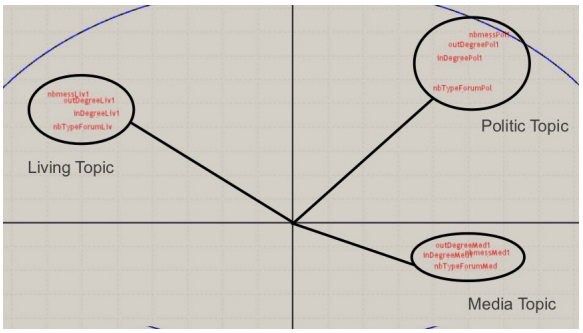


Figure 4. Correlation scatter plot

of *contextual celebrities* for each sub-community.

IV. Conclusion and future works

This paper presents a new exploratory approach to understand on-lines communities based on its sub-communities and give good clues to comprehend the *contextual celebrity* in these sub-communities. A lot of people interact on news websites, this media became increasingly widespread, the dimensionality of data makes it difficult to comprehend. We use the Principal Component Analysis (PCA) to understand how people interact starting from the hypothesis that people focus their participation in one or a very few topics (i.e., context) and not in the website as a whole. The PCA confirmed this hypothesis. This exploratory method finds three kinds of groups defined by the kind of context the users participates in.

The *contextual celebrity*, i.e., a user who participates in one kind of topic and be recognized by his sub-community as a pertinent user, needs to maximize the criteria in one topic. Furthermore, using an enriched social network allows a finer perception of the real interaction between users and brings interesting informations to characterize the community, the sub-communities and the *contextual celebrity* himself.

In perspective, we want to extract the *contextual celebrity* and evaluate the model using a temporal evaluation. We are also interested in the analysis of the users who participate a few (one post) in one topic Who are these people? Why do they participate so little?

References

[1] T. Schoberth, J. Preece, and A. Heinzl, "Online communities: a longitudinal analysis of communication activities," in *Proceedings of the 36th Annual Hawaii International Conference on System Sciences, 2003.* IEEE, pp. 1–10.

[2] D. Hymes, *Foundations in sociolinguistics: An ethnographic approach*. Psychology Press, 2003.

[3] E. Goffman, *The presentation of self in everyday life*. Harmondsworth, 1978.

[4] S. Golder and J. Donath, "Social roles in electronic communities," *Internet Research*, vol. 5, pp. 19–22, 2004.

[5] E. Gleave, H. Welser, T. Lento, and M. Smith, "A conceptual and operational definition of'social role'in online community," in *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*. IEEE, 2009, pp. 1–11.

[6] F. Lorrain and H. White, "Structural equivalence of individuals in social networks," *Social networks: a developing paradigm*, vol. 1, p. 67, 1977.

[7] A. McCallum, X. Wang, and A. Corrada-Emmanuel, "Topic and role discovery in social networks with experiments on enron and academic email," *Journal of Artificial Intelligence Research*, vol. 30, no. 1, pp. 249–272, 2007.

[8] A. Wolfe and D. Jensen, "Playing multiple roles: Discovering overlapping roles in social networks," in *Proceedings of the 21st International Conference on Machine Learning, Workshop on Statistical Relational Learning and its Connections to Other Fields.*, 2004.

[9] J. Zhang, M. Ackerman, and L. Adamic, "Expertise networks in online communities: Structure and algorithms," in *Proceedings of the 16th International conference on World Wide Web*, 2007, pp. 221–230.

[10] N. Agarwal, H. Liu, L. Tang, and P. S. Yu, "Identifying the influential bloggers in a community," in *WSDM '08: Proceedings of the international conference on Web search and web data mining*. New York, NY, USA: ACM, 2008, pp. 207–218.

[11] P. Domingos, "Mining social networks for viral marketing," *IEEE Intelligent Systems*, vol. 20, no. 1, pp. 80–82, 2005.

[12] H. Welser, E. Gleave, D. Fisher, and M. Smith, "Visualizing the signatures of social roles in online discussion groups," *Journal of Social Structure*, vol. 8, no. 2, 2007.

[13] I. Himelboim, E. Gleave, and M. Smith, "Discussion catalysts in online political discussions: Content importers and conversation starters," *Journal of Computer-Mediated Communication*, vol. 14, no. 4, pp. 771–789, 2009.

[14] J. Kelly, D. Fisher, and M. Smith, "Friends, foes, and fringe: norms and structure in political discussion networks," in *Proceedings of the 2006 international conference on Digital government research, May*, 2006, pp. 21–24.

[15] D. Fisher, M. Smith, and H. Welser, "You are who you talk to: Detecting roles in usenet newsgroups," in *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, 2006, pp. 59b–59b.

[16] S. Angeletou, M. Rowe, and H. Alani, "Modelling and analysis of user behaviour in online communities," *The Semantic Web–ISWC 2011*, pp. 35–50, 2011.

[17] J. Chan, E. Daly, and C. Hayes, "Decomposing discussion forums and boards using user roles," in *AAAI Conference on Weblogs and Social Media*, 2010, pp. 215–218.

[18] M. Forestier, A. Stavrianou, J. Velcin, and D. A. Zighed, "Roles in social networks: Methodologies and research issues," *Journal of Web Intelligence and Agent Systems*, p. To appear, 2011.

[19] M. Forestier, J. Velcin, and D. Zighed, "Extracting social networks to understand interaction," *Proceedings of the International Conference on Advances in Social Network Analysis and Mining (ASONAM 2011)*, pp. 213–219, 2011.

[20] J. Donath, "Identity and deception in the virtual community," *Communities in cyberspace*, pp. 29–59, 1999.

[21] K. Pearson, "Principal components analysis," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 6, no. 2, p. 559, 1901.

[22] R. Rakotomalala, "Tanagra: un logiciel gratuit pour lenseignement et la recherche," *Actes de EGC*, vol. 2, no. 3, pp. 697–702, 2005.

 ISBN: 978-1-61208-176-2

Unique Domain-specific Citizen Identification for E-Government Applications

Peter Schartner
Institute of Applied Informatics
System Security Group
Klagenfurt University
9020 Klagenfurt, Austria
Email: peter.schartner@aau.at

***Abstract*—When discussing the security of e-government applications one of the most crucial aspects is the identification of the users (aka citizens). On the one hand, the authorities and the users want to be sure beyond doubt that a certain action or record is related to the correct individual. On the other hand users do not want to have their actions or data in different domains (like health-care, taxes, register of residents, legal authorities) being linked to each other by the authorities. In this paper, we propose an efficient mechanism, which guarantees both, unique identification and inter-domain privacy protection. First of all, the proposed scheme is a replacement for the domain-specific citizen identifier defined by the Austrian authorities, but the scheme may be used as well in other scenarios, depending on unlinkable and unique identifiers.**

***Index Terms*—e-government; system-wide unique identifier; domain-specific identifier; pseudonyms; anonymity; UUIDs; GUIDs.**

I. Introduction

Concerning e-government, accountability of actions or records is one of the most important requirements. On the one hand, authorities would like to know, which user (citizen) has taken a certain actions or, which user is the owner of a certain record. On the other hand, the users do not want their actions or records being mixed up with actions or records of other users. So both groups need and want accountability, which strongly depends on unique identification of the related instances.

Despite the need for unique identification of citizens, most commonly data protection acts (or similar legal requirements) prohibit the (direct) use of unique identifiers (like passport serial numbers or social insurance numbers) outside the scope of these identifiers. Additionally, the users demand privacy protection, i.e., users do not want their actions (or records) to be linked across different domains. For example, data related to health care should not be linkable to data of social insurance and vice versa. So for both reasons, legal regulations and privacy protection, we need some sort of digital pseudonym, which uniquely identifies a citizen, but hampers the linking across domains.

The remainder of this paper is structured as follows. First we will briefly discuss related work concerning the generation of unlinkable (and unique) identifiers. After analyzing the drawbacks of the different schemes, we introduce the so called concept of collision-free numbers, which are used to generate system-wide unique domain-specific citizen identifiers. The paper will close with some modifications of the proposed scheme and open problems, which are the scope of future research.

II. Related Work

In this section, we will briefly discuss internet/industrial standards and some straightforward techniques for the generation of unlinkable unique identifiers. Besides these, we will discuss the approach of the Austrian authorities in more detail, as flaws in this approach brought up the idea of designing a replacement. Basically, all generation processes described, "try" to provide two properties for the identifiers at the same time:

- **Uniqueness:** No two (or more) citizens should be assigned the same identifier. If this happens, this could result in records or actions of different persons becoming inseparably mixed up.
- **Privacy:** Identifiers used in different domains should not be linkable to each other. In some scenarios even the linking between the person and its identifier should be impossible, which results in complete anonymity. In principle this results in the requirement that identifiers "should look" random.

A. UUIDs and GUIDs

A widely adopted approach for system-wide unique system parameters are **universally unique identifiers** (UUIDs, see [1]) and **globally unique identifiers** (GUIDs, see [2]), Microsoft's implementation of UUIDs). There exist several variants of GUIDs, but these variants either use the MAC address to guarantee uniqueness or they employ hash-functions or purely pseudo-random values. Except the first one, which violates the privacy requirement (the MAC address may be linkable to the user), none of them can guarantee uniqueness (since cryptographic hash-functions always come with the risk of duplicates).

B. National Citizen Identifier

In Austria, each individual is assigned a unique so called base number (B – Basiszahl), which is either the individual's

 ISBN: 978-1-61208-176-2

number in the central register of residents, or B is the number in the so called supplementary register, if the person is not subject to registration. Since the Austrian data protection act prohibits the direct use of the base number B, the derivation scheme for unique unlinkable domain-specific identifiers consists of two major phases (see Figure 1):

1) Disguising the base number B by use of an injective transformation, which results in the so called base identifier (bID).
2) Deriving the domain-specific citizen identifier ($dcID$) by use of the base identifier (bID) and the domain identifier (dID).

Phase 1: Disguising the base number consists of the following steps:

1) Input: base number B (12 decimal digits)
2) Binary encoding of B (5 byte)
3) Extension of B to fill two 3DES blocks (16 byte = 128 bit) by use of the following format:

$$b = B \,||\, seed \,||\, B \,||\, B,$$

where $||$ denotes the concatenation of bit strings and $seed$ is a secret constant (8 bit), only known by the authority, which holds the register of residents.
4) Encryption of the binary representation of b by use of 3DES [3] in CBC mode [4], [5] (no padding needed since the input is a multiple of the block size):

$$c = \text{3DES}_k(b),$$

where the secret key k is only known by the authority, which holds the register of residents.
5) For the ease of further usage, the result is Base64-encoded [6] to form the base identifier:

$$bID = \text{Base64}(c).$$

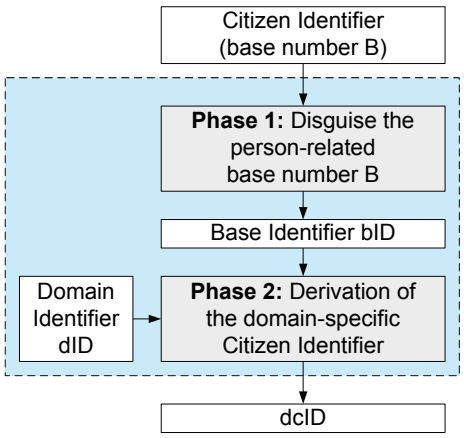


Fig. 1. Original derivation of the domain specific citizen identifier dcID

Analysis: The system-wide unique base number B is encrypted by 3DES (a block cipher) using a fixed key and seed. Hence this is an injective function and the output, the base identifier bDI, is system-wide unique as well. From the security point of view it has to be mentioned, that in case the secret key k becomes publicly known, all base identifiers can be decrypted and actions identified by use of the base identifier can be linked to persons by use of the base number B. Additionally, each individual is assigned exactly one base identifier. Hence, actions or records identified by use of the base identifier may be unlinkable to persons directly, but ar at least linkable to each other. If one of the linked actions or records provides information about its initiator or holder, all other linked actions or data sets can be linked to this specific person.

To overcome the problem of inter-domain linking discussed above, the Austrian authorities proposed to use a derivation scheme, which generates a so called domain-specific citizen identifier ($dcID$) based on the individual's base identifier (bID) and a domain identifier (dID). In order to avoid duplicates, the domain-specific citizen identifiers should be unique with high probability.

Phase 2: The derivation of the domain-specific citizen identifier $dcID$ from the base identifier bID and the domain identifier dID consists of the following steps:

1) Input: Base identifier bID (Base64-encoded) and domain identifier dID (according to the corresponding regulation [7] two to fife ISO/IEC 8859-1 [8] upper case characters)
2) Concatenation ($||$) of base identifier bID, a fixed prefix and the domain identifier dID to form the string s:

$$s = (bID \,||\, \text{“}+\text{”} \,||\, URN - prefix \,||\, dID),$$

where $URN - prefix$ is the ISO/IEC 8859-1 string “urn:publicid:gv.at:cdid+”.
3) Calculation of the SHA-1 hash [9] of s, which results in a 160 bit value h:

$$h = \text{SHA-1}(s)$$

4) Finally, h (as a binary string) may be directly used as domain-specific citizen identifier $dcID$ or may be Base64-encoded before transmission or printout.

Analysis: Since domain-specific citizen identifier are derived by the use of a hash-function, there is the risk of duplicates regardless the fact, that the input to the hash-functions are system-wide unique. Hence there is the risk of inseparable records of different individuals e.g., in E-Government databases.

C. Other Approaches

There exist at least three straight forward solutions for generating random *and* system-wide unique parameters:

- **Centralized generation and check** obviously avoids duplicates but is quite inefficient concerning storage (all previously generated parameters have to be stored for later comparison) and communication (each instance, which needs a parameter has to wait for the centralized generator to send it). Additionally, the centralized generator has full control over the generating process and knows all parameters.

 ISBN: 978-1-61208-176-2

- With **Local generation and (centralized) check**, only the generation itself is done locally, but the comparison against all previously generated parameters has to involve all other generators or a centralized service. Again, efficiency and security are quite questionable.
- **Local generation based on pseudo-random number generators** (PRNG, see [10] for details) can avoid centralized storage and comparison and is efficient in terms of memory and communications. But in order to avoid duplicates, all PRNGs have to use a common key or common secret parameters. So, if one of them is compromised, all of them become insecure. Additionally, the generated parameters are no longer random, but pseudo-random and this approach is not suitable for software implementation, because by use of software, the system-wide key (or secret parameter) cannot be protected sufficiently.

A more sophisticated approach is the so called **location- and time-based generation**, which simply uses location and time provided by a GPS receiver to derive a unique seed for the generation process. The idea behind this concept: two generation processes cannot take place at the same place *and* the same time. Besides the fact that the GPS signal will not be available at all locations, the according paper does not specify, how (pseudo-) randomness and uniqueness are maintained (see [11] for details).

D. Summary of Related Work

Summarizing the related work, we see that none of them fulfills both requirements at the same time: system-wide uniqueness and privacy protection (full or inter-domain unlinkability).

III. Preliminaries

After briefly revisiting basic cryptographic algorithms used in this paper, we will present the core building block of unique domain-specific identifiers: so called collision free number generators (CFNG, introduced in [12], [13]).

A. Cryptography

We assume that the reader is familiar with **Symmetric Encryption** (like DES [14], 3DES [3], or AES [15]) and **Hash-functions** (SHA-1 [9] or RIPEMD160 [16]), and refer to [10] for further details.

In order to keep the output of symmetric encryption as short as possible, we will employ **Ciphertext Stealing**. Let l_B be the block-length of a symmetric encryption function E. Let u be a plaintext, where $l_B < l_u \leq 2l_B$. If u is encrypted straightforwardly by padding u up to $2l_B$ bits and then encrypting two blocks, the length of the corresponding ciphertext c is $l_c = 2l_B$. Using the CBC mode [4], [5] with ciphertext stealing [17], c can be generated such that $l_c = l_u$. This works as follows: First u is cut into the blocks u_1 and u_2, where $l_{u_1} = l_B$ and $l_{u_2} = l_u - l_B$. Then u_1 is encrypted by use of E and a properly chosen key k resulting in a block $c_1||c_2$, where $l_{c_1} = l_u - l_B$ and $l_{c_2} = l_B - l_{c_1}$. Then the block $c_2||u_2$ is encrypted by use of E and the same key k resulting in the block c_3. This works, since $l_{c_2} + l_{u_2} = l_B$. The ciphertext of u is then $c_1||c_3$ and contains sufficient information to compute u, if k is available. The length of c is $l_c = l_{c_1} + l_{c_3} = l_u$. An example for a 64 bit block cipher (like DES) encrypting an 79 bit input can be found in Figure 2.

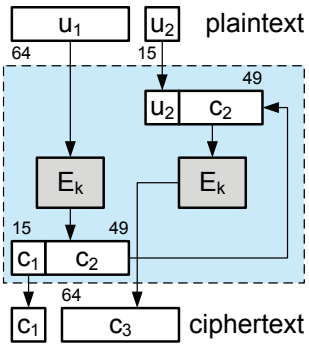


Fig. 2. CBC mode with ciphertext stealing

Details on **Elliptic Curve Cryptography (ECC)** can be found in [18]). For the ease of reading this paper we will just define the basics of ECC.

Definition: Let $E(\mathbb{Z}_p)$ be an elliptic curve group, where p is an odd prime. Let $P \in E(\mathbb{Z}_p)$ be a point of prime order q, where $q|\#E(\mathbb{Z}_p)$. The *Elliptic Curve Discrete Logarithm Problem (ECDLP)* is the following: Given a (random) point $Q \in \langle P \rangle$ and P, find $k \in \mathbb{Z}_q$ such that $Q = kP$.

By $SM(k, P)$ we henceforth denote the **Scalar Multiplication** kP in $E(\mathbb{Z}_p)$. It is believed that the ECDLP using $l_p \approx l_q \approx 160$ is secure against powerful attacks like Pollard's rho algorithm [18].

Point Compression [19]: A point on an elliptic curve consists of two coordinates and so requires $2l_p$ bits of space. It is clear that for every x-value there exist at most two possible y-values. Since they only differ in the algebraic sign, it suffices to store only one bit instead of the whole y-value. A point (x, y) can hence be stored as $x||b$, where $b = y$ MOD 2, and then only requires $l_p + 1$ bits of space.

This has the only drawback that if we want to include this point in some computations, we first have to compute the two possible y-values and then decide by b, which of them is correct. In our case, we are only interested in saving space. There is no necessity to compute y here.

B. Collision-free Number Generators

In [12], we proposed so called collision-free number generators (CFNGs) as a mechanism for generating random but system-wide unique (cryptographic) parameters. Basically, these generators disguise a unique (eventually publicly known) parameter by use of a randomizer. In the scope of e-government identifiers, the information being disguised will be the digital identity of the citizen. The resulting parameter will be a system-wide unique domain-specific digital pseudonym.

 ISBN: 978-1-61208-176-2

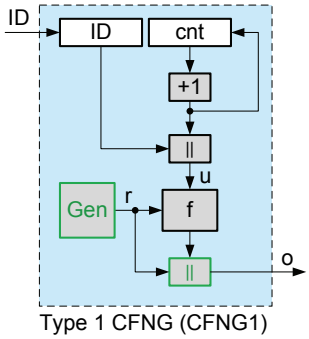


Fig. 3. Basic construction of Collision-free Number Generators (CFNGs)

The output o of a basic – type 1 – CFNG (denoted as CFNG 1 in the remainder of this article, also see Figure 3) is of the form

$$o = f(u, r)||r = f_r(u)||r = \text{CFNG1}(),$$

with f being an injective mixing transformation for an arbitrary but fixed randomizer r and u, r defined as above. We suggest to either use an injective one-way mixing-transformation for f_r according to Shannon [20] (e.g., symmetric encryption) or an injective probabilistic one-way function, based on an intractable problem (e.g., the discrete logarithm problem [10]).

In this paper, we will just revisit the proofs of uniqueness. For a detailed discussion of randomness, efficiency and privacy protection, we refer the reader to [12].

Theorem: *Outputs of Type 1 CFNGs are unique during their lifetime.*

Proof: Consider two outputs of two arbitrary type 1 CFNGs: $o_1 = \text{CFNG1}_1() = f_{r_1}(u_1)||r_1$ and $o_1 = \text{CFNG1}_2() = f_{r_2}(u_2)||r_2$, with r_1, r_2 being random and $u_1 = ID_1||cnt_1$ and $u_2 = ID_2||cnt_2$. With respect to the randomizers r_1 and r_2, there are two cases:

1) $r_1 \neq r_2$: This directly means that $o_1 \neq o_2$.
2) $r_1 = r_2 = r$: Now, both calls of the generators employ the same randomizer and f_r becomes injective. Hence $f_r(u_1)$ and $f_r(u_2)$ will be different if and only if $u_1 = ID_1||cnt_1$ and $u_2 = ID_2||cnt_2$ differ in at least one bit. This is always true, because
   a) different generators use different identifiers ($ID_1 \neq ID_2$), and
   b) if we call the same generator twice (i.e., $ID_1 = ID_2$), the values cnt_1 and cnt_2 will differ, because the counter is incremented at each call of the generator.

   Hence the outputs o_1 and o_2 will be different again.

□

When analyzing CFNGs, which employ a block cipher E (CBC mode with ciphertext stealing) for f ($o = \text{E}_r(ID||cnt)||r = c||r$), it is obvious that the identity of the generator is not protected sufficiently. Everybody who gets hold of an output o can retrieve the identifier ID of the according generator by simply decrypting c by use of r: $ID||cnt = \text{D}_r(c)$.

We will see that this may not be a problem in certain application scenarios; but, in order to guarantee the protection of the generators ID we have either to change our requirements on f, or we have to slightly change the design of CFNGs.

- To provide privacy, f has to be a cryptographic one-way function. Candidates include injective probabilistic one-way functions based on an intractable problem like the (ECC) discrete logarithm problem [10].
- In the case that f is a (bijective) symmetric encryption function, we can employ an additional (injective) one-way-function g to the output or to the randomizer of the original CFNG, which results in the variants depicted in Figure 4 (CFNG 2 and CFNG 3):
  1) The first variant simply hides the output of a type 1 CFNG by use of function g:

$$o = g(\text{CFNG1}()) = \text{CFNG2}(),$$

  2) The second variant only hides parameter r (which is needed to invert function f) by use of function g:

$$o = f(u, r)||g(r) = f_r(u)||g(r) = \text{CFNG3}(),$$

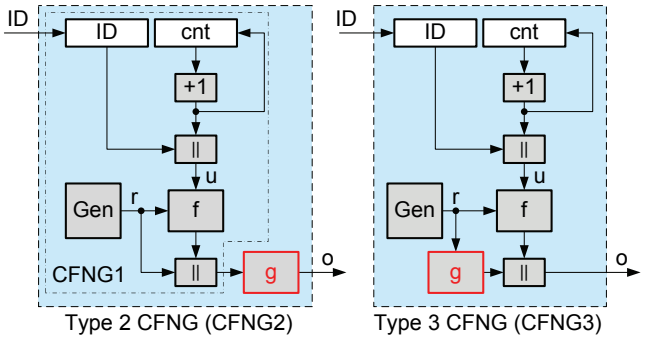


Fig. 4. Variants of Collision-free Number Generators

Corollary: *Outputs of type 2 and type 3 CFNGs are unique during their lifetime.*

Proof: Function g, applied to the unique inputs in type 2 and type 3 CFNGs is a injective one-way-function. Hence g applying g cannot destroy the uniqueness of the outputs.

□

IV. Our Proposal: Unique Domain-Specific Citizen Identifiers

In this section we present three methods to generate unique and unlinkable domain-specific Citizen Identifier:

- **Method 1** (the basic principle) may be directly used as a replacement for the scheme described in Section II-B as it uses the same inputs (and inputs lengths) and generates outputs of equal length.
- **Method 2** uses slightly different (shorter) inputs, but employs more randomness to disguise the inputs. Nevertheless it may also be used as a replacement for the old citizen identifiers.

 ISBN: 978-1-61208-176-2

- **Method 3** uses the base number (60 bit) as the source of uniqueness instead of the base identifier (128 bit) as methods 1 and 2 do. As with method 2, shorter inputs to the encryption function allow more randomness.

A. Basic Principle

Based on a type 2 CFNG employing elliptic curve cryptography (ECC – see [18] for details), elliptic curve scalar multiplication (SM) and point compression (PC) we will now present a generator for system-wide unique and inter-domain unlinkable identifiers. As in the original scheme, our replacement (see Figure 5) generates 160 bit identifiers. But in contrast to the original scheme, these outputs are provably system-wide unique, as we employ type 2 CFNGs (see Figure 4 left) parameterized as follows:

1) Inputs: Base identifier bID (128 bit) and domain identifier dID (five uppercase letters encoded in 24 bit).
2) Starting from the output length of 160 bit we have to subtract one bit to encode the y-coordinate of the ECC-point, 24 bit to encode the domain identifier and 128 bit to store the base identifier. This results in 7 bits remaining for the randomizer.
3) The unique and inter-domain unlinkable identifiers $dcID$ is of the following form:

$$dcID = \text{PC}(\text{SM}((\text{DES}(u, k) \,||\, r), P)),$$

where where P is a so-called generator point of the elliptic curve, $|r| = 7$ and $k = \text{msb}_{56}(H(r))$. In order to reduce redundancy and the bit length of the input of the encryption function, we omit the constant URN-prefix.

Since we employ DES to encrypt the base identifier, we need to expand the randomizer r (7 bit) to 56 bit. This can easily be achieved by use of a hash-function H (e.g., RIPEMD160 [16] or SHA-1 [9]) and a trimming function msb, which extracts the 56 most significant bits: $k = \text{msb}_{56}(\text{H}(r))$. Note that the low entropy of key k is not a severe problem here, because the only purpose of k (based on randomizer r) is to hamper brute force attacks (by a factor of $2^7 = 128$ in this setting).

B. Variant 1

Up to now, the Austrian e-government act [21] and the corresponding domain regulation [7] define just 35 different domain identifiers (see Table I).

So spending 24 bit to store the domain identifier dID is a massive overhead. A more practical solution is reducing the bit length of dID by half (i.e., to 12 bit) and using a binary encoding instead of the text encoding. By this, the length of the randomizer r can be enlarged by 12 bit, which results in $|r| = 19$ bit.

C. Variant 2

This variant directly uses the base number B (5 byte = 40 bit) instead of the base identifier bID (128 bit) and hence shortens the input of the encryption function by 88 bit. We will use some of the bits to embed additional data X, which might hold a counter in order to provide different identifiers within the same domain. The remainder of the bits will be used to enlarge the randomizer to 80 bit and replace DES with SKIPJACK [22] (block length 64 bit and key length 80 bit) in CBC mode with ciphertext stealing. This finally results in unique and unlinkable domain-specific identifiers of the form:

$$dcID = \text{PC}(\text{SM}((\text{SKIPJACK}_r(B \,||\, X \,||\, dID) \,||\, r), P)),$$

with $|B| = 40$ bit, $|X| = 15$ bit, $|dID| = 24$ bit and $|r| = 80$ bit.

Note that the direct use of the base number (which can also be the passport or social insurance number) may be prohibited by law. In this case, variant 1 or the basic scheme have to be used.

TABLE I
CURRENT LIST OF DOMAIN IDENTIFIERS

e-government domain regulation – appendix to § 3 – part 1					
AR	AS	BF	BW	EA	EF
GH	GS	GS-RE	JR	KL	KU
LF	LV	RT	SA	SF	SO
SO-VR	SR-RG	SV	UW	VT	VV
WT	ZP				
e-government domain regulation – appendix to § 3 – part 2					
BR	HR	KI	OI	PV	RD
VS	VS-RG	ZU			

V. CONCLUSION AND FUTURE WORK

We are aware of the fact that the proposed scheme is first of all a replacement for a national standard for generating unlinkable domain-specific identifiers (which does not completely fulfil its own requirements). But nevertheless, provably system-wide unique unlinkable and domain specific identifiers based on collision-free number generators (CFNGs), parameterized as defined in Section IV-A, may be employed in other application scenarios as well. These scenarios include identifiers in the context of e-business, the replacement of UUIDs and GUIDs [12], temporary MACs for untraceable network devices [23], and digital pseudonyms [24], [25].

REFERENCES

[1] P. Leach, M. Mealling, and R. Salz, "RFC 4122 – A Universally Unique IDentifier (UUID) URN Namespace," 2005, (retrieved: 12/2011). [Online]. Available: http://www.ietf.org/rfc/rfc4122.txt
[2] Microsoft Developer Network, "Globally Unique Identifiers (GUIDs)," http://msdn.microsoft.com/en-us/library/cc246025.aspx, 2008, (retrieved: 12/2011).
[3] National Institute of Standards and Technology (NIST), "FIPS Special Publication 800-67: Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher," 2008.
[4] ISO/IEC, "ISO/IEC 10116: Modes of Operation of an n-bit Block Cipher," ISO/IEC, 1991.
[5] National Institute of Standards and Technology (NIST), "FIPS Special Publication 800-38A: Recommendation for Block Cipher Modes of Operation – Methods and Techniques," 2001.
[6] S. Josefsson, "RFC 4648 – The Base16, Base32, and Base64 Data Encodings," 2006, (retrieved: 12/2011). [Online]. Available: http://www.ietf.org/rfc/rfc4648.txt

 ISBN: 978-1-61208-176-2

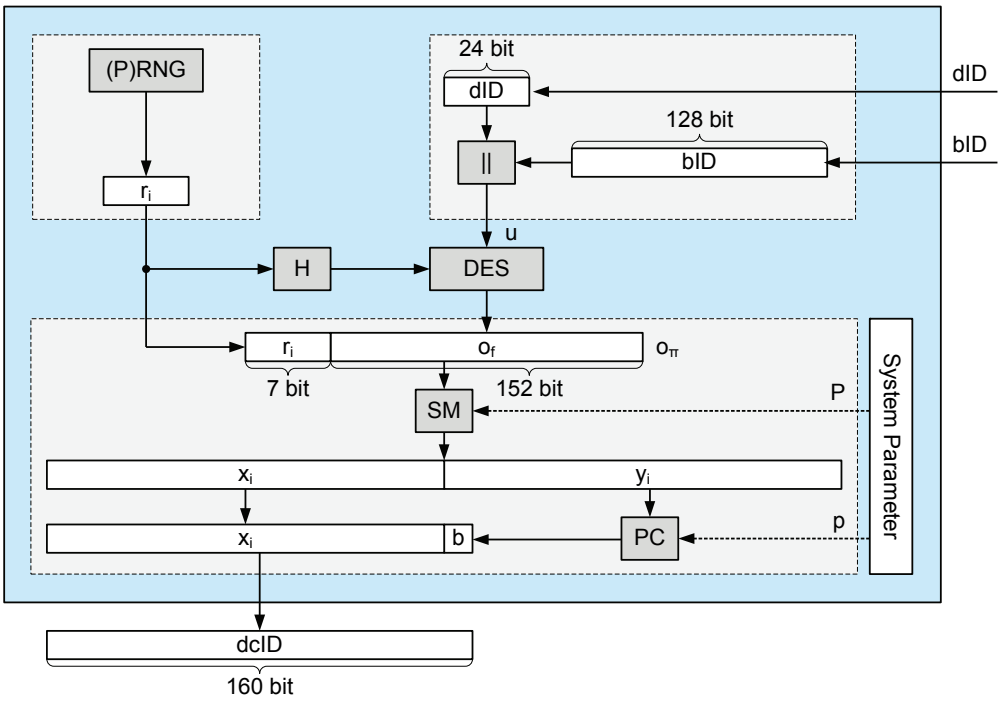


Fig. 5. New implementation of the domain specific citizen identifier $dcID$

[7] Republik Österreich, "Verordnung des Bundeskanzlers, mit der staatliche Tätigkeitsbereiche für Zwecke der Identifikation in E-Government-Kommunikationen abgegrenzt werden (E-Government-Bereichsabgrenzungsverordnung – E-Gov-BerAbgrV) StF: BGBl. II Nr. 289/2004, (Fassung vom 14.9.2011)," 2004, (retrieved: 12/2011). [Online]. Available: http://www.ris.bka.gv.at

[8] ISO/IEC, "ISO/IEC 8859-1:1998, Information technology – 8-bit single-byte coded graphic character sets – Part 1: Latin alphabet No. 1," ISO/IEC, 1998.

[9] National Institute of Standards and Technology (NIST), "FIPS Publication 180-2: Secure Hash Standard," 2002.

[10] A. Menezes, S. Vanstone, and P. V. Oorschot, *Handbook of Applied Cryptography*. CRC Press, Inc., 1996.

[11] IPCOM, "Method of generating unique quasi-random numbers as a function of time and space. PriorArtDatbase, IPCOM#000007118D," 2002, http://priorartdatabase.com/IPCOM/000007118 (retrieved: 12/2011).

[12] M. Schaffer, P. Schartner, and S. Rass, "Universally Unique Identifiers: How To Ensure Uniqueness While Protecting The Issuer's Privacy," in *Security and Management*, S. Aissi and H. Arabnia, Eds. CSREA Press, 2007, pp. 198–204.

[13] P. Schartner, "Random but system-wide unique unlinkable parameters," *JIS – Journal of Information Security*, vol. 3, no. 1, January 2012, ISSN Print: 2153-1234, ISSN Online: 2153-1242, in print. [Online]. Available: http://www.scirp.org/journal/jis

[14] National Institute of Standards and Technology (NIST), "FIPS Publication 46-3: Data Encryption Standard (DES)," 1999.

[15] ——, "FIPS Publication 197 – Advanced Encryption Standard (AES)," 2001.

[16] H. Dobbertin, A. Bosselaers, and B. Preneel, "Ripemd-160: A strengthened version of ripemd," in *Proceedings of Fast Software Encryption (FSE)*, ser. LNCS, D. Gollmann, Ed., vol. 1039. Springer, 1996, pp. 71–82.

[17] C. Meyer and S. Matyas, *Cryptography: A New Dimension in Computer Data Security*. John Wiley & Sons Inc, 1982.

[18] D. Hankerson, A. J. Menezes, and S. A. Vanstone, *Guide to Elliptic Curve Cryptography*. Springer-Verlag, 2004.

[19] IEEE, "Std 1363-2000: IEEE Standard Specifications for Public-Key Cryptography," 2000.

[20] C. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28(4), pp. 656–715, 1949.

[21] Republik Österreich, "Bundesgesetz über Regelungen zur Erleichterung des elektronischen Verkehrs mit öffentlichen Stellen (E-Government-Gesetz – E-GovG), BGBl. I 10/2004, (Fassung vom 3.3.2011)," 2010, (retrieved: 12/2011). [Online]. Available: http://www.ris.bka.gv.at

[22] National Institute of Standards and Technology (NIST), "SKIPJACK and KEA Algorithm Specifications, ver. 2, 29," 1998.

[23] M. Schaffer and P. Schartner, "Untraceable Network Devices," Klagenfurt University (Austria) – System Security Group (syssec), Tech. Rep. TR-syssec-06-04, November 2006.

[24] P. Schartner and M. Schaffer, "Unique User-Generated Digital Pseudonyms," in *MMM-ACNS*, ser. Lecture Notes in Computer Science, V. Gorodetsky, I. Kotenko, and V. Skormin, Eds., vol. 3685. Springer, 2005, pp. 194–205.

[25] ——, "Efficient privacy-enhancing techniques for medical databases," in *BIOSTEC (Selected Papers)*, ser. Communications in Computer and Information Science, A.L.N.Fred, J. Filipe, and H. Gamboa, Eds., vol. 25. Springer, 2008, pp. 467–478.

 ISBN: 978-1-61208-176-2

Sociological Reflections on E-government

Maria João Simões
Department of Sociology
University of Beira Interior, Covilhã
Researcher at Research Centre of Social Sciences (CICS), University of Minho, Braga
Portugal
e-mail: mariajoaosimoes@sapo.pt

Abstract — **The objective of this paper is to present dimensions of sociological analysis that allow a more comprehensive and interpretative analysis of e-government. This effort will contribute to a more critical analysis of its implementation, chosen devices and assessment. The analytical dimensions presented are: (i) citizenship models; (ii) metatheoretical frameworks on society and technology; (iii) the concept of e-government and its articulated domains. It intends to demonstrate that the choice between options of each dimension contributes for different kinds of e-government and results. The e-government is not a neutral issue. The citizenship model adopted, in a very incisive way, makes all the difference in the conception, design, working and results of e-government. The theoretical framework that is underlying to e-government shapes also its design, working and results. But the devices chosen per se are insufficient to characterize an e-government, as their potentialities can be used in a completely different way by people and rulers. Research and projects on e-government are principally focused in e-administration, underestimate e-democracy and e-society that have been analysed in a separate way, which makes difficult a more comprehensive and all-encompassing analysis and assessment of e-government.**

Keywords – e-government; participation; technology; society.

I. INTRODUCTION

Most research projects have a strong descriptive approach, probably because, on the one hand, they have, to a large extent, a practically oriented approach focusing in development projects, applications or case studies. On the other hand, most researches come from the information systems field where the major focus is the conception, design and application of devices. So, it can be said that e-government is still an under-analysed area, from a theoretical and conceptual point of view, as referred by Simões [1], Heeks and Bailur [2], and Lindblad-Gidlund and Axelsson [3].

E-government lacks deepening of theoretical and conceptual frameworks from the social sciences, particularly sociology, which can better explain, in a more comprehensive, interpretative and all-encompassing way, what e-government is, why, what for and how it is implemented. Such frameworks would allow a critical analysis of different visions on e-government, the purposes of its creation in each social context, the interests that underlie its creation, the adopted models of e-government and applications, and also to better understand why different social and technological results are achieved.

What we say above allows us to state that e-government is clearly an interdisciplinary area; more intensive interdisciplinary research is crucial especially between researchers both from information systems and social sciences, particularly sociology of science and technology, political sociology and sociology of organizations. Surprisingly, although political sociology is a widespread field, sociologists have underestimated research on e-government.

This presentation shares, thus, the challenge of Lindblad-Gidlund and Axelsson [4] that argues to be necessary to establish vessels among different scientific areas for rigorous and relevant e-government research.

In this way, based on literature review regarding different theories on the relationship between society and technology, a critical reading of crucial literature on the subject, namely Oliver and Sanders [5], Mayer-Schönberger and Lazer [6], Cunningham and Cunningham [7], based on our experience in projects on local e-government [8] and even on e-participation [9][10], our purpose in this paper is to present critical dimensions, within a sociological point of view, that can allow a more comprehensive and critical approach on e-government research, implementation and assessment.

The dimensions of the critical analysis focused from a sociological perspective are stated in a triptych presentation. Firstly, introducing two ideal types of citizenship models. Secondly, debating different metatheoretical frameworks on society and technology. Thirdly, discussing the concept of e-government and its articulated domains: e-administration, e-democracy and e-society. As a conclusion, final considerations will be presented.

II. E-GOVERNMENT IS NOT A SEPARATE ISSUE OF CITIZENSHIP

First of all, government is one of the most important components of a state: it is the way how it was organized and how rulers establish its interaction with people that we can say if we are dealing, for example, with a dictatorial or democratic state. In this sense, as government is a polysemic term, thus e-government is also polysemic.

 ISBN: 978-1-61208-176-2

But the same happens with a democratic government which is not also a neutral term. The history of democracy was undergone by maximalist and minimalist versions of citizenship. When we talk on e-government, what citizenship version are we talking about? So, we affirm that in any research project, either more theoretical or more empirical, or even in any project of implementation, we have to explain which conception of citizenship is used.

A more active or passive concept of citizenship will induce significant variations on the type of services, its contents, on quantity, quality and kind of available information, and on communication patterns, that is, on the kind of adopted e-government and its working. In that sense it is important to reflect on the different impacts that these different conceptions of citizenship have in e-government and also on chosen applications, as we will discuss further on.

Taking in account the weberian methodology of ideal type, two opposite kinds of political participation are presented [11], constructed for clarification purposes, knowing that there are other models between two ideal types where it can be found different combinations of both.

The passive citizenship is embedded into a liberal perspective, which inspires western democracies and where the citizen role has an individualist and instrumentalist approach, the citizen being granted full rights. The individual has, as Oldfield [12] sustains, not only epistemological priority, but also an ontological and moral one.

For the author, citizenship is seen as a legal status which must be sought, and sustained when accomplished. The state and other institutions are looked in an utilitarian way. It is only expected that they allow the conditions for individuals to maximize their own benefits and reach their goals, without any notion of common welfare present. Though, citizens are demanded to follow a certain set of civic obligations towards the state, namely, to vote, to pay taxes and to defend the country, in the case of external menace.

To liberals, politics is a realm of the government, considered only as what politicians, specialists, political parties and bureaucrats do [13][14].

Although political communication consists of emission and reception, verbalization and listening, liberal theory values the speech and neglects the listening part. It is easier for those in charge to speak rather than to listen.

Participation is largely reduced to choose between several options, thus giving the winner the power to establish the direction of the world we live in. Vote is included in a negotiation model within which choices are predetermined, thus limiting not only the choices opportunities but also the imagination. As Barber [15] says, there are few other possibilities that allow the voters to express their opinions, leaving the citizen as a simple spectator.

In an active participation model, the citizen is a member of a political community, in which he/she assumes a central position. Citizenship is not just a *status*; participation is an objective by itself. In this political *praxis*, being a non participant means, in many ways, that he or she can be an individual but not a citizen [16][17].

For the last author, in order to people being engaged in citizenship practice three conditions are requested all of them necessary but none alone sufficient: resources, participation opportunities and motivation. In the resources domain, beyond the assurance of civil, political and social rights, the economic and social resources (a reasonable living income, education, health, among others) as well as competences regarding the political activity are also crucial.

On the other hand, the participation opportunities have to be assured, which implies the creation and widening of an appropriate institutional setting at several levels (local, national, global and also at horizontal and specialized level) that stimulate the civic participation in general and, in a particular way, a rational understanding and better information of public issues, the participation in agenda setting, deliberation and decision making, among other activities.

Individuals also have to be encouraged to participate, to execute their political rights and duties, that is to say, to be citizens. One cannot expect, as Oldfield [18] writes, that citizenship *praxis* and civic conscience to appear spontaneously. As Steinko [19] states, mobilization implies that people feel there is a link between their daily life, in all spheres, being them local (namely the education, employment, environment issues), national or global.

Actually, considering the growing distance between rulers and people, with the option of these procedures, (e-) government, especially at local or regional context, can become a setting not only to a closer interaction between both but also to reduce the citizens' scepticism concerning politics.

In this model, there is a broader conception of politics, involving all public issues in which the citizens have the right to be involved; «politics describes the realm of *we*» [20].

The access to information is indispensable for the practice of citizenship, but it is only a sufficient condition. Equally important is the kind of information that is delivered. Information has to focus on problems faced by citizens, it has to be contextualized, justified and it should explain the consequences of the political choices that can be made. But the removal of information barriers is not enough [21][22][23].

Speech is equally valued as listening, a recurring and permanent interaction, established upwardly and downwardly, between rulers and citizens.

In this sense, Hacker's [24] political interactivity model has heuristic value. As daily interaction can be simplified, just including a message and its answer, and even get another message from the first user, political interaction requires two additional interactions, as seen in the Figure 1.

The first message (m1) comes from the citizen towards the politician, who, in return, sends his/her feedback to the citizen (m2). The content of this message will determine what happens after the established interaction. In order to reply (or not) to the requested information or the citizens' expectations, citizens can answer back (m3), and the government can answer through political action (m4) or an

 ISBN: 978-1-61208-176-2

explanation (m5) explaining why such course of action can not be fulfilled.

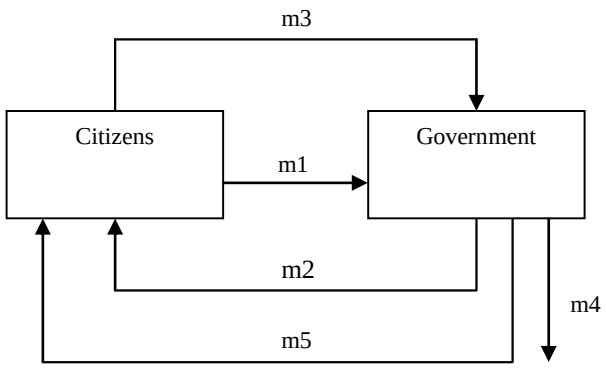


Figure 1 – A basic model of political interactivity [25]

More messages can be exchanged, but this five step flux of interaction model is the basic political interactivity model from which more complex models can be built, emphasizing upward and downward communication between rulers and citizens.

Besides vertical communication, the horizontal kind is also considered to be crucial. On the one hand, the political choice includes deliberation, because individuals, when involved in collective participation, do not always agree on their civic and political concerns. On the other hand, the deliberation help them to overtake their narrow interests; it is through the debate that individuals frequently rejoin themselves, re-evaluate and can reformulate values, beliefs and opinions based upon which they engage in their political participation (Barber [26]; Yankelovich [27]; Oldfield [28]).

According the citizenship model adopted the e-government conception, the design, the implementation and the results and yet the assessment process will be different. Consequently, the services, the kind of information delivered, the communication patterns and the applications will be different. So, the choice of one of these citizenship models makes all the difference from the analytical and empirical point of view and for the achieved results.

III. THEORETICAL FRAMEWORKS ON TECHNOLOGY AND SOCIETY

Chosen the citizenship model to use, we are facing with different theoretical frameworks whenever we engage in further research or when we intend to present and to implement an e-government project. The metatheoretical framework selected has to be clarified because it has consequences on the chosen e-government models, on its implementation and also on achieved results and their assessment.

Many e-government projects and its implementation are based on technological determinism, where the underlying idea is that the properties of technology, namely those used by e-governments, are transposed and absorbed by societies, producing the same effects upon them. This notion forgets that the technological devices are not neutral, since the results will depend on: the specific social and organizational context in which e-government will be implemented; the values and the interests of promoters that influence the choice of e-government models; leadership; the degree of commitment of the staff and the strategy followed; the resistance or involvement on its implementation and yet the way how people and rulers appropriate and shape the devices for their use.

The fact that technological determinism has been dominating this discussion is one of several reasons for the deficit on a more comprehensive and interpretative analysis of e-government.

The option for that metheoretical framework implies that the e-government projects and their implementation, as well as another projects focused on infra-structures, hardware and software and the assessment indicators are to a large extent, including in European Union (EU), predominantly technological [29].

As an alternative to technological determinism, a sophisticated model of analysis on the relation between society and technology can be used, in which technology and society are mutually related, that Simões [30] nominate reciprocal conditioning. This metatheoretical framework has more heuristic potential because it takes in account crucial social aspects (as power, interest groups, conflict, values and so on) that are present in the conception, implementation and execution phases, and therefore in the outputs reached by e-government. It is into this framework that it can be said "e-government is more about government than about 'e'" [31]. On the other hand, it does not underestimate the fact that each technological device can condition our action in a specific direction and not in any other.

In this sense, in such metatheoretical framework, the conception of e-government projects, its implementation and the back office, process, output and demand indicators embraces social and technological aspects.

Contradicting the technological deterministic authors and several designers, the applications choice is not sufficient to characterize an e-government.

Firstly, they can think or install, for example, applications to a horizontal communication (from the more "traditional" as *fora* to the more recent as web 2.0: facebook, twitter and so on), but the rulers or the people, depending of their interests and goals, can make a unexpected use of them. As an example, political parties and rulers in several countries use facebook to communicate with people being interdict the possibility of reply [32]. So, applications designed to a horizontal communication can be used to a vertical and downward one.

Secondly, when a communication device is available, communication might not be started, whether because rulers consider themselves the legitimate representatives of the citizens, whereas these should confine themselves to the episodic election of those, or because citizens are in apathy or do not believe that it is worthwhile, that is, that nothing will come out of their participation. During the timeframe of the Digital Cities Program, the Portuguese Operational

 ISBN: 978-1-61208-176-2

Program for Information Society (POSI) and the Operational Program for Knowledge Society, programs which endured from 1998 to 2006, cities and administrative regions submitted projects to turn themselves into digital cities and territories, e-government being one of the major focus. The great concern with technological modernization and the prevailing technological deterministic perspective lead the promoters to focus mainly on technological infrastructures and software. Most projects encompass devices, although different from one another, allowing horizontal and vertical communication.

We did not find differences neither in the chosen devices in municipalities ruled by either leftist or right parties, nor in the concerns about the actual use of these devices. We present only an exception: in one municipality, where the mayor invested in a more active citizen participation, facing the apathy of people, the mayor said he would focus mainly in face to face participation modalities; only later would he take into consideration information and communication technologies [33]. Nowadays, in some Portuguese cities, new experiences on e-government based in higher citizen participation have to be researched.

Thirdly, there can be some stimulus for citizens to participate, even if there is not any concern from the government with its citizens' worries and anxieties. This is just an illusion of participation, which can be amplified by an automatic answer by e-mail where the citizen participation is thanked, even if there is not a real intention of actually answering and there is a vague promise of taking the citizen participation into account.

Some researchers have yet «observed that the same information system in different organizational contexts leads to different results. Indeed, the same system might produce beneficial effects in one setting and negative effects in a different setting» [34]. Once again, we can point out that technological deterministic authors underestimate several social factors that make the difference in the results of e-government.

IV. E-GOVERNMENT: CONCEPT AND ITS DOMAINS

In the concept of e-government underlies a normative and evaluative component. From a sociological point of view, it is important to understand if there is a political or normative position or rather a scientific one. For example, it is often said that “e-government is better government” [35]. From a scientific point of view, only through empirical evidence can we verify if the e-government can or cannot foster a greater engagement with citizens and enables or not better quality services and policy results.

Several authors, as St-Amant [36], referred three inter-related domains of e-government: e-administration, e-democracy and e-society. The first stands on the administrative modernization issue, on efficiency and efficacy of services and whether electronic services do or do not improve services to citizens, being these principally seen as customers.

In the domain of e-democracy it is debated to which extent ICT can enhance or not the citizen participation and the relationship between rulers and ruled ones. The debate on e-democracy is wide but has been, by large extent, carried out disconnected from e-government. On the other hand, we can face more pessimistic points of view, as Sunstein's [37], or very optimistic ones, as Rheingold's [38], or even more realistic perspectives, stated namely by Simões [39], that identify new opportunities but also new constrains on e-democracy. Either way, this discussion is not the focus of this paper. Regardless of these perspectives having different empirical implications and results, we have different models of political participation in real or virtual context. The chosen model of participation within e-government implies different devices, different uses, different ways of implementation and different achieved results. This is one of the central issues of this paper.

In the domain of e-society it is attempted to verify if the ICT contributes or not to the strengthening of relations between government and civil society organizations, namely NGO, trade unions, universities, I&D institutions, cultural associations, sport clubs and also corporations.

These domains have been frequently studied separately as they were completely different issues. We state that although a research or a project can focus more in one of e-government domains, it has to take into account all them, because they are all closely interconnected as we have emphasized along the paper.

V. FINAL REFLECTIONS

The objective of this paper was to present sociological dimensions of analysis that allow a more comprehensive and interpretative analysis of e-government, its implementation, chosen applications and its assessment.

The dimensions analysed allow a more critical and deeper debate about the interconnection between social and technological factors concerning e-government. Thus, we point out to a more intensive interdisciplinary among different scientific areas for a relevant and rigorous e-government research.

E-government is not a neutral issue. A more active or passive conception of citizenship have significant implications on e-government conception, design, implementation and results.

According to the chosen participation model we will find differences regarding the kind of information and services delivered, the patterns of communication, the intensity and frequency of the interaction between rulers and people.

The adoption of a technological deterministic or a reciprocal conditioning perspective between technology and society have also different implications in e-government, leading to different kinds of e-governments and necessarily different applications. As users can shape applications according to their interests and necessities, the chosen applications *per se* are insufficient to denominate the kind of e-government. Such is only possible with an on-going assessment and with indicators embracing technological and social aspects.

 ISBN: 978-1-61208-176-2

E-government research is, in a large extent, centred in e-administration, it underestimates the e-democracy and e-society, domains largely analysed apart. Although efficiency and efficacy of services are crucial for e-government working, e-government is not a corporation. E-government is more related with people government, with e-democracy and e-society. So, even if a research or a project focuses more on a unique e-government domain, it has to take all of them into account, as they are all closely interconnected. If we do not head towards this path we are drifting apart of the essence of the e-government concept.

Further researches could point to deepen this theoretical reflection on e-government connecting it to more extended empirical research and identifying assessment indicators of e-government that encompass social and technological aspects.

REFERENCES

[1] M. J. Simões, Política e Tecnologia – Tecnologias da Informação e da Comunicação e Participação Política em Portugal, Oeiras: Celta, 2005.

[2] R. Heeks and S. B. Bailur, "Analyzing e-government research: Perspectives, philosophies, theories, methods, and practice", Government Information Quarterly, vol. 24, 2007, pp. 243-265.

[3] K. Lindblad- Gidlund and K. Axelsson, "Communicating Vessels for Relevant and Rigorous eGovernment Research", P. Cunningham and M. Cunningham (Eds.), Collaboration and the Knowledge Economy – Issues, Applications, Case Studies, vol.5, Part 1, 2008, pp. 255-261.

[4] K. Lindblad-Gidlund and K. Axelsson, "Communicating Vessels for Relevant and Rigorous eGovernment Research", P. Cunningham and M. Cunningham (Eds.), Collaboration and the Knowledge Economy – Issues, Applications, Case Studies, vol.5, Part 1, 2008, pp. 255-261.

[5] L. Oliver and L. Sanders, (Eds.), E-Government Reconsidered: Renewal of Governance for the Knowledge Age, Regina: University of Regina, 2004.

[6] V. Mayer-Schönberger and D. Lazer, (Eds.), From Electronic Government to Information Government, Cambridge: The MIT Press, 2007.

[7] P. Cunningham, and M. Cunningham, (Eds.), Collaboration and the Knowledge Economy – Issues, Applications, Case *Studies*, Amsterdam: IOS Press, vol. 5, part 1, 2008.

[8] M. J. Simões (coord.), Dos Projectos às Regiões Digitais - Que Desafios?, Lisbon: Celta, 2008.

[9] M. J. Simões, Política e Tecnologia – Tecnologias da Informação e da Comunicação e Participação Política em Portugal, Oeiras: Celta, 2005.

[10] M. J. Simões, A. Barriga and N. A. Jerónimo, Brave New World? Political participation and new media, SOTICS 2011: The First International Conference on Social Eco-Informatics, pp. 55-60, Copyright (c) IARIA, ISBN: 978-1-61208-163-2, available at http://www.thinkmind.org/index.php?view=article&articleid=sotics_2011_3_10_30040.

[11] Barber (1984), Oldfield (1998) and Held (1996) are crucial authors for our elaboration of citizenship typology; this typology is further discussed in Simões (2005) and in M. J. Simões and E. Araújo (2009).

[12] A. Oldfield, Citizenship and Community: Civil Republicanism and the Modern World, London: Routledge, 1998, (2ª Ed.).

[13] B. Barber, Strong Democracy: Participatory Politics for a New Age, Berkeley, University of California Press, 1984.

[14] D. Held, Models of Democracy, Cambridge: Polity Press, 1996, (2ª Ed.).

[15] B. Barber, Strong Democracy: Participatory Politics for a New Age, Berkeley, University of California Press, 1984.

[16] B. Barber, Strong Democracy: Participatory Politics for a New Age, Berkeley, University of California Press, 1984.

[17] A. Oldfield, Citizenship and Community: Civil Republicanism and the Modern World, London: Routledge, 1998, (2ª Ed.).

[18] A. Oldfield, Citizenship and Community: Civil Republicanism and the Modern World, London: Routledge, 1998, (2ª Ed.).

[19] A. Steinko, "Herramientas para un chequeo de la dinámica democrática", REIS , 1, 1994, pp. 9-35.

[20] B. Barber, Strong Democracy: Participatory Politics for a New Age, Berkeley, University of California Press, 1984.

[21] B. Barber, Strong Democracy: Participatory Politics for a New Age, Berkeley, University of California Press, 1984.

[22] D. Yankelovich, Coming to Public Judgement – Making Democracy Work in a Complex World, Syracuse: Syracuse University Press, 1991.

[23] M. Hale, J. Musso and C. Weare, "Developing digital democracy: evidence from Californian municipal web pages" in Barry Hague and Brian Loader (Eds.), Digital Democracy: Discourse and Decision Making in the Information Age, London, Routledge, 1999, pp. 96-115.

[24] K. Hacker, "The White House Computer-mediated Communication (CMC) System and Political Interactivity" in K. Hacker and J. Dijk (Eds), Digital Democracy, London : Sage, 2000, pp. 105-129.

[25] K. Hacker, "The White House Computer-mediated Communication (CMC) System and Political Interactivity" in K. Hacker and J. Dijk (Eds), Digital Democracy, London : Sage, 2000, pp. 105-129, p. 123.

[26] B. Barber, Strong Democracy: Participatory Politics for a New Age, Berkeley, University of California Press, 1984, p.123.

[27] D. Yankelovich, Coming to Public Judgement – Making Democracy Work in a Complex World, Syracuse: Syracuse University Press, 1991.

[28] A. Oldfield, Citizenship and Community: Civil Republicanism and the Modern World, London: Routledge, 1998, (2ª Ed.).

[29] M. J. Simões and E. Araújo, "A sociological look at e-Democracy", in Patrizia Bitonti and Vanessa Carrieri (Eds.) e-Gov 2.0: pave the way for e-participation, Roma: EuroSpace S.r.l., 2009, pp. 155-161.

[30] M. J. Simões, Política e Tecnologia – Tecnologias da Informação e da Comunicação e Participação Política em Portugal, Oeiras: Celta, 2005.

[31] OECD, The e-government imperative: main findings, www.oecd.org/publications/POL_brief, 07-09-2007, 2003, pp.1-7, p.1.

[32] M. J. Simões, A. Barriga, N. A. Jerónimo, Brave New World? Political participation and new media, SOTICS 2011: The First International Conference on Social Eco-Informatics, pp. 55-60, Copyright (c) IARIA, ISBN: 978-1-61208-163-2, available at http://www.thinkmind.org/index.php?view=article&articleid=sotics_2011_3_10_30040.

[33] M. J. Simões (coord.), Dos Projectos às Regiões Digitais - Que Desafios?, Lisbon: Celta, 2008.

[34] J. Fountain, "Central Issues in the Political Development of the Virtual State", The Network Society and the Knowledge Economy: Portugal in the Global Context Conference, March 4-5, http://www.umass.edu/digitalcenter/research/pdfs/jf_portugal2005_centralissues.pdf, 26-09-2007, 2005, pp.1-29, pp. 4-5.

[35] OECD, The e-government imperative: main findings, www.oecd.org/publications/POL_brief, 07- 09 -2007, 2003, pp. 1-7, p.1.

[36] G. St-Amant, "E-gouvernment: cadre d'évolution de l'administration électronique", Revue Management et Système d'Information, vol. 10, nº1, ABI/INFORM Global, 2005, pp. 15-39.

[37] C. Sunstein, Republic.com, Princeton: Princeton University Press, 2001.

[38] H. Rheingold, Electronic Democracy Toolkit, available at www.well.com/user/hlr/electrondemoc.html, 06-09-2000, 1996.

[39] M. J. Simões, Política e Tecnologia – Tecnologias da Informação e da Comunicação e Participação Política em Portugal, Oeiras: Celta, 2005.

SSEDIC: Building a Thematic Network for European eID

Victoriano Giralt
Central ICT Services
University of Málaga
Málaga, Spain
e-mail: victoriano@uma.es

Hugo Kerschot
IS Practice
Brussels, Belgium
e-mail: hugo.kerschot@is-practice.eu

Jon Shamah
EJ Consultants
Harrow, United Kingdom
e-mail: jshamah@ejconsultants.co.uk

***Abstract*—Digital Identity is a critical element for a digital society as proposed by the Digital Agenda for Europe. The width and breadth of the subject makes it mean different things in different sectors, even to different projects funded by the European Commission. Thus, having a network that provides a platform for all the stakeholders of electronic identity to work together and collaborate to prepare the agenda for a proposed Single European Digital Identity Community, is of prime importance to the achievement of said goals. The network, SSEDIC, is working on identifying the actions and the timetable for the Digital Agenda and the successful launch of the European Large Scale Action and European Innovation Partnerships, as well as to provide a multi stakeholder planning resource to assist its implementation. A first batch of deliverables will be presented to the European Commission at the end of February 2012 and then made available to the public. This paper will present the SSEDIC expert network as it is now, what has been done to build the network and the accomplishments of its first year. But, the most important aim of this paper is to increase awareness about SSEDIC and reach out to some valuable contributors that had not yet been identified to get them involved in the network.**

Keywords-Electronic Identity; Single European Digital Community; Digital Agenda for Europe.

I. INTRODUCTION

Every European digital [1] (by 2020) this is the ambitious goal set by Commissioner Neelie Kroes for the Digital Agenda for Europe (DAE)[2]. In order to achieve this goal, a single European digital community is needed, and the DAE (Digital Agenda for Europe) sixteen key actions [3] show how it could be achieved. Also, the DAE calls for stakeholder involvement to reach the goals.

Key action 16 in the Digital Agenda for Europe [3] proposes a Single Digital Identity Community and scoping that is the purpose of SSEDIC. The network has built a platform where stakeholders can identify the actions and the timetable for their resolutions to result in the successful launch of the European Large Scale Action and European Innovation Partnerships (ELSA/EIP). This cannot happen out of thin air, thus SSEDIC [8] is building upon the ELSA/EIP thematic consultations carried out by the ELSA/EIP eIDM (*E*lectronic *ID*entity *M*anagement) Expert Working Group [4] and today's Large Scale Pilots (LSPs) such as, but not limited to, STORK [5], PEPPOL [6] or SPOCS [7].

The SSEDIC thematic network [8], during its first year in existence, has established a series of stakeholder groups in sectors outlined in the ELSA/EIP report [4]. Each of the groups will consider, through further consultations, the political, economic, social, technical, legal and environmental aspects of a single European digital community.

This network is built gathering experts from 35 partners and an initial group of associated partners. The former provide 67 experts in electronic identity (eID) who are picked for their knowledge of the eID or stakeholder domain rather than just as representatives of organisations. This has been a fundamental criterion for partners to ensure that the views and consultations are of the highest value and relevance to this highly important thematic. The later can grow as much as wanted and one of the main aims of the present paper is to increase the visibility of SSEDIC in as much pertinent communities as possible, because the ambition of this network is to build a community of high level European and international experts up to 2013 and, if possible, beyond.

A stakeholder is defined as any group or individual that can affect or is affected by the achievement of SSEDIC [8] objectives. They often have differing interests and may put conflicting pressures on the project. The consortium needs to attend to a rich variety of claims and interests of stakeholder groups in the industry, yet at all times needs to profile a coherent identity of the project to each and every one of these groups. A wide range of persons and groups exist with differing legitimate interests in SSEDIC. Recognising and addressing their needs and interests will in turn enhance the performance of the project, ensure that it is aligned with market realities, and secure its continued acceptance.

Additionally an overarching and integrated view of the accumulated results and inputs from the various stakeholder sectors will be taken in order to build an overview and impact assessment of a single European digital community on the overall European Community and also on individual EU Member States.

The high level outcome of the SSEDIC thematic network [8] is to provide a wide ranging and valuable consultation-based resource and consensus which will enable the European Commission to understand the **roadmap** that must be addressed within the ELSA/EIP programme to progress Europe's single

 ISBN: 978-1-61208-176-2

digital community vision as outlined by the DAE [2] across each and every sector of the European Community. And this output is intended to be a thought-through and widely agreed blueprint for step-by-step actions which can feed directly into the future ELSA/EIP-programme and drive that programme towards a successful conclusion delivering a European digital identity community.

In order to have real impact on society, the SSEDIC thematic network [8] is not intended to be an academic exercise, it is an action plan with a roadmap for the DAE for the coming decade. As such, SSEDIC results need to reflect a transformational shift in the way everyone in the European community will think, behave, transact and indeed live in the coming years. The vision to be established by SSEDIC can be a beacon to the rest of the world, demonstrating how the efficiency of a digital community can be translated into cultural and economic leadership. Although the vision will need to be technology led, it cannot be technology targeted; rather, the vision should integrate goals derived from stakeholder sector needs and benefits and within a holistic framework of actions.

The present paper will describe the SSEDIC thematic network background and work methodology, that has lead to a first batch of deliverables that will seed the final results expected by the end of 2013.

II. Background

In 2009, the European Large Scale Actions consultations commenced the description of a Single European Digital Community. A number of SSEDIC partners made contributions via the ELSA/EIP eIDM Expert Working Group, which resulted in the *ELSA/EIP report* [4]. This consultation described an interoperable network of independent but regulated Identity Service Providers, many possibly Public-Private partnerships, which would make an eID (not *National* eID) available to each citizen within each Member State, while retaining full freedom of choice for the individual.

SSEDIC partners include representatives of member states with a *National eID* infrastructure as well as countries with alternative eID models.

SSEDIC partners include experts and organisations that have participated in the eID initiatives of Norway, Denmark, Austria, Italy, Belgium and Germany. Also, SSEDIC keeps contact with the eGov subgroup via the Commission, in order to assess eGovernment policy and in particular eID policies of the Member States.

The Large Scale Pilots such as STORK [5], SPOCs [5], epSOS [10], PEPPOL [6] and other projects are critical to the success of the Single European Digital Community. The technology and standards being evolved will form the cornerstones for interoperability, not only for cross-border use cases, but also to establish and cement trustworthy relations between Identity Service Providers in the same countries. Many of these projects could not join SSEDIC as full partners themselves, but are contributing as observers with strong inputs to the consultations, thus providing an opportunity to ensure the continuity and sustainability of their key outputs. On the other hand, many of the contributors to these projects are SSEDIC partners, thus there is strong involvement of the coordinators of the LSPs in the SSEDIC Network and overlap of many members. This will ensure that full mutual benefit is realised and the standards, components and demonstrator experience can be incorporated into the consultative outcomes. As SSEDIC progresses, contacts with other EU projects in all sectors will be fostered.

The Higher Education sector is a special case as they already have electronic identity infrastructures in production across Europe and beyond, that connect research and educational institutions both inside Member States and inside and outside the Union. Three partners of project SEMIRAMIS [11], that deals with eID supporting the movement of students inside the European Higher Education Area [12], are SSEDIC partners. Other partners connect SSEDIC to experts networks in Europe dealing with eID federations in research and education.

STORK [5] is essentially a proof of concept of technical interoperability, but the project has established the basic building blocks of the infrastructure that will ensure eID interoperability at European level, including common code for an architecture and interoperability platform which will be released under EUPL. These building blocks address other dimensions beyond technical interoperability, such as multilateral trust mechanisms, framework for security assessment of national infrastructural components, harmonisation of Quality Authentication Assurance mechanisms, etc. Additionally, the pilots that the project have set-up, and which make use of the above mentioned infrastructure, have a strong potential beyond the project time frame. SSEDIC will use as basis the studies, technology overviews, and prototypes on new and upcoming technologies produced in STORK for the consultations with the experts groups. SSEDIC has a strong link to STORK through six common partners. These strong relationships will result in additional benefits and allow for making suggestions as to how to exploit the achievements of STORK and its pilots into the future.

III. Work Plan and Methods

There is a **communication plan** supported by an **action plan**, that cover the 36 months duration of the project, to use and disseminate knowledge at different levels. The main areas addressed by the plan are:

- promoting the SSEDIC thematic network identity and results within the network and beyond;
- sharing general knowledge, specific information and documents through open source collaborative tools;
- organising meetings, seminars and workshops in different formats for the use of the entire thematic network team, using traditional as well as innovative techniques;
- continuously integrating the SSEDIC thematic network knowledge in partners dissemination channels

A. PESTLE

The aggregation and organisation of the consultation data will be based on PESTLE, which stands for:

 ISBN: 978-1-61208-176-2

- **P**olitical
- **E**conomic
- **S**ociological
- **T**echnological
- **L**egal
- **E**nvironmental

PESTLE analysis is an audit of an organisation's environmental influences with the purpose of using this information to guide strategic decision-making. A PESTLE analysis is a useful tool for understanding the *big picture* of the environment in which any organisation is operating [13].

All consultation work will take into account these six aspects, where relevant.

The PESTLE analysis will allow for the production of time lines with specific actions to be carried out in each of the six aspects.

B. Activities

SSEDIC has established a series of stakeholder groups in sectors contributing to the EIP. Each of the groups considers, through further consultations, the political, economic, social, technical, legal and environmental aspects of a single European digital community. This groups are formed off the experts from the 35 partners plus experts from associated partners, which number should grow by accretion over the 36 months lifespan of the project. Each stakeholder work programme consists of brainstorming workshops, strategy papers and joint meetings with more general sector organisations to gain a fuller understanding of the requirements and prerequisite actions for delivery of the vision in that sector. Hard data will be built to consider the impact and opportunities of the single European digital community in the short, medium and long term.

C. Work Packages

SSEDIC has organised the consultation at 3 levels and each of these levels has a dedicated work package:

1) Stakeholders Sector Consultation
2) Technology and Infrastructure Consultation
3) Business Model and Regulations Consultation

plus three global work packages dedicated to coordination, dissemination and outcome management.

The materials produced by the three consultation work packages will be merged by the outcomes management one and used for dissemination.

The stakeholders sector consultation, due to the large number of stakeholders sectors, has required these to be formed into 6 groupings described in Figure 1.

The consultation on technology and infrastructure is split into natural areas of interest:

- Security
- Privacy and Ethics
- Enrolment
- Identity Models
  - Nonrepudiation
  - Interoperability
  - Identity Service Provision
  - Authentication
- High Level Architecture
  - Standards
  - Integration
  - Resilience
  - EU projects
- Accessibility
  - Credentials
  - Accessibility
- Operations
  - Regulations
  - Monitoring
  - Quality of Service

The Business Model and Regulatory consultation is considering business models, revenue models, and regulatory regimes needed to establish a successful vision. It is looking in depth at the ELSA/EIP Thematic consultation [4] and further expanding these business aspects. It is also looking at Member State issues, interoperability and also cross stakeholder benefits. This work pack age overlaps many stakeholder and technology issues such as privacy, ethics and standards.

D. Tools

The ambition of the SSEDIC thematic network, is to build a community of high level European and international experts. This community is being built via virtual tools such as a dedicated online workspace and online conferences as well as via real live events integrated in the EEMA (European association for eIdentity and security) [9] conference programme and other major European events.

Our strength is the quality of our network individuals, not just their affiliate organisations. Each of the experts is picked for their knowledge of the eID or stakeholder domain rather than just as representatives of organisations. This has been a fundamental criterion for partners to ensure that the views and consultations are of the highest value and relevance to this highly important thematic project.

The network is composed by four groups of partners who check and balance each other in their different sectors:

- Industry: who have strong contacts in the private sector
- Public sector: with access to government and local agencies and stakeholders
- Academic partners: with critical reflections on industry services and public sector requirements and interests such as Erasmus students.
- Small and medium sized consultancies: who have strong influence across all domains.

As much dedication and knowledge as persons can dedicate to the SSEDIC network, there is a clear need for technological tools to support them. Thus, SSEDIC has established a main dissemination web site, http://www.eid-ssedic.eu/ with both public and private areas for document and information sharing

 ISBN: 978-1-61208-176-2

Government	Society	Business	Leisure	Finance	Transport
eParticipation	eHealth	Telecommunications	Media	Banking	Automotive
eJustice	eInclusion	Manufacturing	Culture/Arts	Insurance	Aviation
Law enforcement	Emergency Services	eCommerce	Sports	Intermediation	Shipping
Regulation authorities	Environment	FMCG/Retail	Adult content	Internet payment	
Border control	Education	Food-chain	eGaming	Mobile payment	
Local authorities	NGO's	Agriculture & Fisheries	Social Networks		
	eConsumer				

Figure 1. Stakeholders Sectors Grouping

and download. Final results will be published in this web site when available.

A second technological tool is an online social network and community building one provided by one of the SSEDIC partners, located at http://ssedic.syncsphere.com/. This tool is used for discussions on the consultation process and allows the network experts to share knowledge and opinions.

The third big technological tool is the on-line surveying one, that has allowed the network to carry out a first survey on eID gathering input from 211 experts on the matter.

Of course, other tools such as email or teleconferencing are also being used to coordinate the experts network.

Non technological tools like publications and presentations in relevant events are also being used for the dissemination and outreach of the SSEDIC thematic network.

E. Barriers

The SSEDIC thematic network identified four barriers that could hinder its efforts and prepared mitigation actions:

1) Lack of response from stakeholders
   A large, and increasing, number of experts minimises this risk
2) Non-representative opinions
   Minimised by carefully monitoring each sector
3) Lack of funds for in-depth citizen consultation
   Mitigated by the use of online surveying tools
4) Contrary interests of stakeholders
   Mitigated by already existing consensus for the need of a common vision of eID and the wide involvement of stakeholders in SSEDIC

IV. Project First Year

The project officially started on December 15th 2010 with the SSEDIC kick-off meeting. The main aims of this first year the project have been to:

- Introduce the project to all stakeholders;
- Make stakeholders aware of basic information regarding the SSEDIC project
- Inform stakeholders of the portal as an information resource
- Help promote the project in conferences and other events
- Initiate interaction with stakeholders and receive feedback and reactions about the project that will be used in media relations and in designing the dissemination plan
- Promote participation of institutions and organisations through the Project Forums and surveys in the portal
- Focus from the start on establishing a favourable reputation for the project and consortium
- Profile a coherent identity of the project to each and every one of the stakeholder groups

V. Results

It must be emphasised that SSEDIC is a thematic network for consultation and is not mandated to make decisions on technologies. However it is envisaged that technical recommendations and statements of Best Practice will be agreed and presented among the final outcomes.

A. Expected Final Outcomes

SSEDIC [8] final outcomes should enable the European Commission to instigate measures to allow Member States fulfil the vision.

1 SSEDIC will create for each Stakeholder Sector an electronically retrievable resource, containing the main consultations, consensus and impacts. This should enable the entire European Community to conduct actions that will ultimately contribute to or benefit from the Single European Digital Identity Community.

2 SSEDIC will create, at the technical level, an electronically retrievable road map of critical actions, milestones and time lines. This road map will outline how to achieve the vision of the Single European Digital Identity Community.

3 SSEDIC will create a combined topological mapping of all the Stakeholder and Technical Sectors, which will be integrated into this single high level road map. This mapping will ensure that role/responsibility divisions and expectations are clear to all stakeholders.

4 SSEDIC will create a combined impact assessment summary, across all the Stakeholder and Technical Sectors integrating all business and regulatory issues.

B. Achieved Results

Over its first year of existence, the SSEDIC thematic network has already been able to produce some good quality deliverables with interesting information, relevant. These materials will be available on the main dissemination web site, http://www.eid-ssedic.eu/, once they are presented to the

 ISBN: 978-1-61208-176-2

European Commission. We are not authorised to reveal the contents until this presentation has occurred. The results will be available under a Creative Commons license in order to achieve as high an impact as possible.

Many of the partners have done presentations in sector events as well as in institutional meetings like the Euro Parliament or the International Telecommunications Union, and events organised by network partner EEMA [9].

Each of the relevant work packages have produced a report with information of the activities and consultations that have been carried out in their area, interim reports and conclusions, calls for action and future activities.

211 eID experts all over Europe where surveyed on items about

- The impact (or not) of eID in a professional environment
- Their policy views on different aspects of eID
- The adoption of the eID technology in a business environment
- Possible future business and governance models on eID infrastructure
- Security and privacy aspects of eID

The results of this survey have already been processed and transformed into a report that will also be presented to the European Commission at the end of February 2012 and the made publicly available at the SSEDIC main dissemination web site, http://www.eid-ssedic.eu/.

VI. Conclusions and Future Work

The network start has not been an easy one because it is difficult to coordinate such a big and heterogeneous group of people as busy as field experts. But, the willingness of this same group of people to collaborate towards a common vision of eID in Europe, and beyond, has been key to a successful start.

Work is already progressing as predicted and the first results have seen the light, though not yet, as the writing of the present paper, general availability.

There are clear definitions for further work during the two remaining years of the project into 2013 and ways to improve what has already been done such as:

- More sector reports
- Surveys focused on different sectors
- Increase the level of discussion
- A big eID event involving all interested parties, stakeholders and LSPs

Acknowledgements

The present paper is not in any way personal work, it is just a recount of the work done by the SSEDIC thematic network consortium partners for winning a bid and to set up the project, and the work of all partners, consortium and associate, and other contributors to the consultations who have given excellent input to the results achieved so far. Of course, the authors would like to thank all of these people and also those LSPs and CIPs that have paved the way for our network to be a need.

References

[1] N. Kroes, *Every European Digital*, Neelie Kroes Blog, URL: http://blogs.ec.europa.eu/neelie-kroes/every-european-digital/ retrieved: November 21st, 2011.

[2] European Commission, *Digital Agenda for Europe*, URL: http://ec.europa.eu/information_society/digital-agenda/index_en.htm retrieved: November 21st, 2011.

[3] European Commission, *Digital Agenda for Europe: key initiatives*, URL: http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/10/200&format=HTML&aged=0&language=EN&guiLanguage=en retrieved: November 21st, 2011.

[4] ELSA Thematic Working Group on Electronic Identity Infrastructure, *European Large Scale bridging Action (ELSA/EIP): Electronic Identity Management Infrastructure for trust worthy services*, European Commission, Directorate-General for Information Society and Media, URL: http://ec.europa.eu/information_society/activities/egovernment/docs/studies/elsa_eid_thematic_report_final.pdf retrieved: November 21st, 2011.

[5] STORK project, *Secure idenTity acrOss boRders linKed*, URL: http://www.eid-stork.eu/ retrieved: November 21st, 2011.

[6] PEPPOL project, *Pan-European Public Procurement OnLine*, URL: http://www.peppol.eu/ retrieved: November 21st, 2011.

[7] SPOCS project, *Simple Procedures Online for Cross- Border Services*, URL: http://www.eu-spocs.eu/ retrieved: November 21st, 2011.

[8] SSEDIC project, *SSEDIC: Building a Thematic Network for European eID*, URL: http://www.eid-ssedic.eu/ retrieved: November 21st, 2011.

[9] EEMA, *European Association for eIdentity and Security*, URL: http://www.eema.org/ retrieved: November 21st, 2011.

[10] epSOS project, *Smart Open Services for European Patients*, URL: http://www.epsos.eu/ retrieved: November 21st, 2011.

[11] SEMIRAMIS project, *Secure Management of Information across multiple Stakeholders*, URL: http://www.semiramis-cip.eu/ retrieved: November 21st, 2011.

[12] EHEA, *European Higher Education Area*, URL: http://www.ehea.info/ retrieved: November 21st, 2011.

[13] F.J. Aguilar, *Scanning the business environment*. Macmillan, New York, 1967.

 ISBN: 978-1-61208-176-2

Designing National Identity:

An Organisational Perspective on Requirements for National Identity Management Systems

Adrian Rahaman, Martina Angela Sasse
Computer Science Department
University College London
London, United Kingdom
a.sallehabrahaman@cs.ucl.ac.uk, a.sasse@cs.ucl.ac.uk

***Abstract* - Many National Identity Management Systems today are designed and implemented with little debate of the technologies and information required to fulfil their goals. This paper presents a theoretical framework detailing the organisational requirements that governments should consider to implement effective identity systems. Analysis is based on publicly available documentation on the implementation of National Identity Systems in the countries of Brunei, India, and the United Kingdom. The findings and the framework highlight the importance of clearly defining the purpose of the system, which has implications on the authenticity, uniqueness, and uses of identity; failure to consider these components is likely to lead to ineffective identity systems and policies.**

Keywords – identity; identity mangaement system; organisation; government; policy

I. INTRODUCTION

Identity is a valuable resource that shapes and defines social interactions [1] by reducing uncertainty and building trust between parties. Governments have traditionally provided identities for their citizens, and used them to manage the provision of services. In an age of growing travel and migration, and facing threats such as illegal immigration, crime and terrorism, *"many governments today are now trying to reassess their identity policies in light of technological changes"* [2].

Governments have tended to view National Identity Management Systems (N-IDMSs) technology as a silver bullet – or at least cornerstone – to tackling these problems, but fail to consider the complexity of delivering such systems [3]. In the UK, attempts to short-cut debate and deliver a system quickly led to adoption of a system that has now been scrapped [4]. Without proper consideration of purpose and operational requirements of the N-IDMS, it is unlikely they will deliver their stated goals.

Convinced that requirements for a strong proof of identity means an increase in security, governments have not examined the use of identity beyond it. But personalised and customer/citizen-centric services mean that identity is no longer just a mechanism for individuals to access resources - it has itself become a valuable resource being accessed by organisations to inform their decisions [5-7].

Still, most research on this topic focuses on identity as a security mechanism. For example, a very comprehensive model for governments' transition to digital N-IDMSs [8] mainly describes its use for online authentication purposes; [9] developed an IDMS architecture that places identity as a layer below information resources.

The research presented in this paper moves beyond the security perspective, viewing identity as a strategic resource. The aim of the study was to uncover organisational identity requirements, and their effects on the design and implementation of IDMSs (The term organisation as used within this document refers to the organisation that is implementing the IDMS).

Section II below explains the methodology followed in this study. In Section III, we present our findings, and explore the processes **of identity construction** and **identity use**. Section IV highlights the importance of **purpose**, which then ties all the findings into a single framework. In Section V, we discuss the implications for future N-IDMSs: to meet their defined purpose, the key factors of **authenticity, uniqueness,** and the objectives of the **relying parties**have to be clearly defined.

II. METHODOLOGY

Our research used a case study approach - a systematic analysis of the identity phenomenon in 3 different cases [10] of N-IDMS implementations in Brunei, India, and the United Kingdom.

The data analysed on the UK and India N-IDMSs was publicly available system documentation published by the respective lead agencies (IPS and UIDIAI respectively); for the Brunei case study, interview sessions with key government officials were recorded and transcribed; interviews were conducted with:

- 3 employees from the lead agency (BruNIR) that deal with strategy and implementation of the system.
- 2 employees from a security organisation that works with the lead agency on the N-IDMS.
- 3 employees from a Relying Party that makes use of the N-IDMS as an authenticator
- 1 employee from a Relying Party that was seen as a prime candidate during the initial phases but is now considering launching its own IDMS system.

The data was analysed using Grounded Theory, a method to develop theory that is grounded in data [11]; i.e., it does not start with a preconceived theory, but seeks to generate new theory through a systematic collection and analysis of data [12].

 ISBN: 978-1-61208-176-2

III. RESULTS

Our analysis revealed that organisational identity requirements, and its eventual impacts on the final design of the system, can be divided into two main areas; **identity creation** and **identity application.**

A. Identity creation

When an identity system is first implemented, a new and unique context is created, within which identities need to be instantiated. It is within this newly created context that an organisation needs to ensure the *correctness* of identities being enrolled. This process is important because it affects the integrity of the identity, and has an impact on the type and amount of personal information being collected and stored.

The challenge of the enrolment process it is that it involves the verification of unknown individuals. Organizations typically fall back on two main criteria when enrolling new identities: **authenticity** and **uniqueness.**

1) Authenticity

Authenticity describes the truthfulness of an identity created within the IDMS. It seeks to answer the question, *is the individual who he says he is?* Organisations typically ensure authenticity of an individual's identity by verifying his/her biographical information against different sources. Organisations can vary the source of information by choosing between two different schemes:

- **Introducer-based schemes** build on the concept of personal referrals - having an already enroled individual vouch for the authenticity of the individual who is attempting to enrol in the system.
- **Document-based schemes** are designed around the use of available identification documents provided by other organisations (bank statements, utility bills, etc). Such schemes rely on third-party organisations confirming the authenticity of enrolling individuals.

While an organisation can choose between the two sources of information, it is limited by the context of its implementation; the main contextual factors that influence the applicability of these schemes are **universality** and **intimacy**.

a) *Universality*

This concept captures how many members of the target population already possess widely accepted forms of identity documents. These are identities that individuals have established with third-party organisations with whom they have a trust relationship; examples of such organisations include banks, utilities, and municipalities that an individual has interacted with for a period of time.

The degree of universality in the target population will affect an organisation's ability to rely on a document-based scheme for authenticity. Specifically, having little to no universality would remove this option, because many individuals would not be able to provide the required documents.

The case study of the Indian NIDMS provides an example of the problem arising from low universality. A large section of the population has been locked out of both public and private services; the weak identity infrastructure has resulted in a fragmented approach to the enrolment in current systems, placing large burdens on most of the poor population to prove themselves, and being denied access to basic services as a result.

"...every time an individual tries to access a benefit or service, they must undergo a full cycle of identity verification. Different service providers also often have different requirements in the documents they demand, the forms that require filling out, and the information they collect on the individual. Such duplication of effort and identity silos increases overall costs of identification, and cause extreme inconvenience to the individual. This approach is especially unfair to India's poor and underprivileged residents, who usually lack identity documentation, and find it difficult to meet the costs of multiple verification processes." [13]

Given the aim is to provide access to its poorer citizens, India cannot create an N-IDMS that relies on a document-based scheme. Therefore, the UIDAI has chosen an introducer-based scheme, *"where introducers authorized by the Registrar authenticate the identity and address of the resident*" [14].

In contrast, the abandoned UK N-IDMS was strongly motivated by prevention of criminal activities and illegal immigration. While the system documentation does state that the UK N-IDMS will make it easier to prove identity [15], UK citizens were not being denied services because of a lack of identity - most of the population had recognised forms of identity provided by third-party organisations.

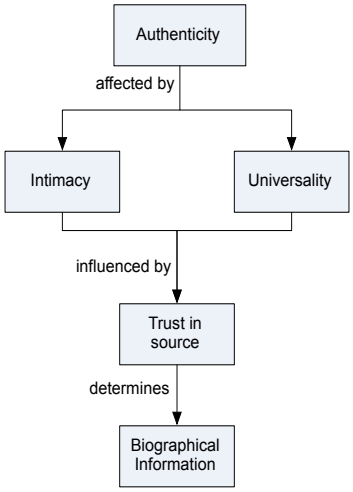


Figure 1. Organisatons Identity Creation process - Authenticity

The UK system took a document-based approach, requiring individuals who enrol for an identity to provide several different documents as proof for the authenticity of the claimed identity [16]. The government required that individuals provide documents that have some form of unique identifier such as passport numbers, driving license numbers, national insurance numbers and *"any number of any designated document, which is held by him"* [17]. This

 ISBN: 978-1-61208-176-2

creates an *information net* around the claimed identity, which the government can then use to ensure authenticity by verifying the individual's personal information with the relevant third party organisations.

b) *Intimacy*

The concept of intimacy captures how much of the targeted population is already known to the organisation. High levels of intimacy imply that the organisation can have more confidence in an introducer-based scheme, because it can support a transitive trust arrangement that extends from known individuals to unknown ones.

The effects of intimacy can be seen in the Bruneian context and its combined approach to ensuring authenticity, incorporating elements of both a document-based and introducer-based scheme. Running an identity system since 1949 [18], the government has been enrolling identities of all individuals born and staying within the country, and thus have established a great deal of intimacy with its population. While individuals are required to provide their birth certificates as proof during enrolment, the government also records the identity numbers of the individual's parents. This in effect creates a hybrid document-introducer-based scheme where the authenticity of the individual is being proven with a minimal amount of documentary evidence, which is further supported by linkages to introducers that are already enrolled within the system.

While India has an introducer-based scheme, the government's choice in the matter is forced by unsatisfactory levels of universality. However, India now faces the problem that there is not enough intimacy to support introducers, as used in the Brunei case. Having never registered identities of past populations, the UIDAI in India cannot currently rely on parents as introducers to the system. Therefore, the government has devised a scheme to artificially boost intimacy through a set of defined trusted recognised introducers.

Introducer and document-based schemes are not orthogonal. Both make use of transitive trust to ensure the authenticity of the claimed identity. The document-based scheme is basically an institutionalised version of the introducer-based scheme. At the centre of the document-based scheme is the reliance on identity documents that have been produced by third-party institutions, which fulfil the role of introducer. In the end, the authenticity of the claimed identity is verified by a trusted third party.

2) Uniqueness

Apart from authenticity, organisations also need to consider uniqueness - that is to ensure that identity cannot be enrolled more than once into the identity database. Organisations' desire for uniqueness is driven by concerns of identity fraud, where individuals might attempt to enrol multiple times, potentially using multiple personas, to gain extra benefits. Organisations typically attempt to tackle this issue of *de-duplication* through the use of biometric data [19].

Today, organisations can choose between various different biometric solutions; facial, fingerprint, and iris recognition being current solutions of choice. Organisation's choice of biometric are influenced by 3 main criteria; **obligations, performance, and population**.

a) *Obligations*

The first hurdle an organisation faces when choosing a biometric technology are the obligations that it must conform to, such as **international standards and current practices.**

International standards influence the choice of biometrics, especially if individuals' identity is meant to be portable across different countries, organisations, or contexts. If, the organisation aims to achieve interoperability, this determines not only the type of biometric used, but also the format in which the data is stored. For example, the UK government defended its choice of fingerprints with the need to comply with standards published by International Civil Aviation Organisation (ICAO) [20]; however, the ICAO standards only proscribe *how* fingerprints should be implemented *if* they are used on such documents – but do not proscribe the use of fingerprints itself [21].

Similarly, although the UIDAI did not focus on ensuring compatibility with other countries, adhering to an accepted standard remained an issue, to help create a consistent and portable identity within India's large borders. The report from the Indian Biometric Committee recommended the implementation of biometrics based on international standards (ISO 19794), stating that the *"standards are widely accepted, and best embody previous experiences of the US and Europe with biometrics"* [19].

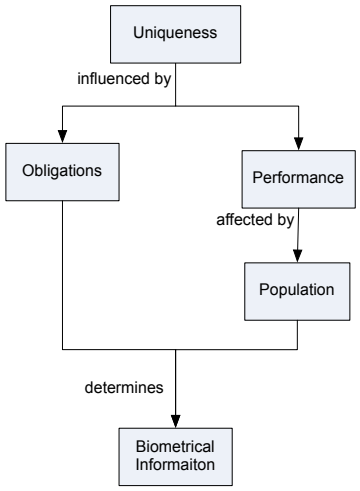


Figure 2. Organisatons Identity Creation process - Uniqueness

Organisations also face obligations around current practices that it, or other organisations that it might work with, have already implemented. The existence of current practices around the use of certain biometrics implies the availability of experience, expertise, and infrastructure around that particular biometric. Having such familiarity with a particular biometric can help to ease the

 ISBN: 978-1-61208-176-2

implementation of a new identity system that makes use of the same biometric.

In the UK context, this can be seen in the relationship between the Identity and Passport Service (IPS) and the Immigration and National Directorate (IND) [20]. Prior to the plans for an N-IDMS, the IND had already been processing, recording, and storing facial and fingerprint biometrics of foreigners for the purpose of UK visa applications. Thus, when the IPS finalised its plans for the N-IDMS it chose to build on IND's systems, directly storing fingerprints and facial biometrics on IND databases. In the Bruneian context, the biometrics deployed in the previous N-IDMS was carried forward into the new, making use of fingerprints and facial photographs that they were already familiar with.

b) *Performance*

Aside from its obligations, organisations are also influenced by the performance of the various biometrics; these can be expressed in terms of **accuracy** and **human interpretation**.

Accuracy captures the ability of the biometric technology to correctly match biometrics presented for verification against biometric templates that have been previously recorded. During enrolment, organisations typically want to prevent individuals from enrolling more than once. This is achieved by choosing biometrics that provides the required levels of accuracy. Failure to match comes in form of False Acceptance - an impostor being wrongly accepted against an enroled identity - and False Rejection, an enroled individual being rejected by the system [22] (Further discussion of these measures is outside this scope of this work). Organisations should also consider the ease of which the biometric can be circumvented. For example, Facial Biometrics is *"considered a poor biometric for use in de-duplication"*, as an individual can easily avoid identification through *"the use of a disguise, which will cause False Negatives in a screening"* [19].

While the use of biometrics to ensure uniqueness is typically an automated process, a manual form of checking identity is required when a false rejection is encountered. Since the system is unable to accurately distinguish between two or more biometrics, some form of backup authentication is required to confirm or deny the false rejection. Therefore, having a biometric that enables quick manual checking becomes a necessity. Most biometric do not lend themselves easily to manual inspection. As a result, facial biometrics becomes attractive to organisations simply because it provides a backup option through **human interpretation** [19].

"We use AFIS, Automated Fingerprint Identification System. All the fingerprints captured will be processed with the fingerprint matching, and this is very useful when the citizen does registration of the card. This is to ensure that one citizen holds one card and number only. Those who register will go through the AFIS matching, and if it is OK, then we will do the registration. Otherwise there will be human intervention; a matching process, the system will list the possible candidates that match, but normally we go for a 100% match. There is a possibility of 70, 80, 90 and 100% match by fingerprints. The system also makes use of facial image, from the entries identified by AFIS. So it's easy for us to do the matching, we can even assign the matching tasks to the clerk, by looking at the facial image and the percentage. It is very straight forward and user friendly." [23]

c) *Population*

An organisation's performance considerations are in turn mediated by the population characteristics in 3 ways: **size, compatibility, and geographic diversity**.

First of all, organisations need to consider the **size** of the target population. Large population sizes can negatively affect the accuracy of the biometric. The choice of the 10-finger biometrics proposed in the UK and Indian scheme was made on those grounds. The Indian Biometric committee [19] established that *"False Acceptance Rate is linearly proportional to gallery size"*; using a 2 fingerprint scheme with a population size of 1.2 billion, the FAR was estimated to be 14%, which is well above the 1% mark that they required. Therefore, it was recommended to proceed with a 10-fingerprint scheme, which was estimated to provide a 0% FAR, maintaining the uniqueness of individuals in the database.

The second population characteristic is **compatibility**, which captures the suitability of the biometric for use on the targeted population. Compatability is commonly captured by tests demonstrating that accuracy is not affected by characteristics of the target population (e.g. skin tone); the lack of such studies was highlighted by the Indian Biometric committee [19].

However, organisations must also consider other real world cultural compatability factors that are not captured by these tests. For example, the Indian Biometric Committee highlighted the use of Lawsonia Interims (Henna) by women, stating that it can prevent the accurate collection of fingerprints as *"sensors may not properly capture fingerprint features."* Another example is the large percentage of population in India who are *"employed in manual labour"*, and thus provide *"poor biometric samples"*, as their fingerprints have been worn away by the nature of their work. Because of these issues, iris is now seen to provide a better match for this population [19]. In Brunei, the BruNIR has encountered problems with the compatibility of fingerprints:

"…only one, the taking of the fingerprint. Because they can get worn out, and those are very difficult to capture. We identified that since the beginning of the project, and we came up with a solution to make use of moisturizer. It helps, but that is the major problem" [23].

Geographic diversity deals with the dispersion of the targeted population across the nation. This can affect the accuracy of the biometric because of varying conditions under which the biometric data is collected and used, and because procedures may be used differently. When a population is spread across a large geographic area, the organisation is unlikely to be able to collect all the information on its own; it will likely adopt an accredited enrolment strategy, where authorised third parties collect

 ISBN: 978-1-61208-176-2

information on their behalf. UK and India are prime examples of third-party enrolment, using private organisations to enrol and capture individual biometrics. This can result in *"several non-technical factors that can impact accuracy more significantly than technical accuracy improvement efforts, such as the lack of adherence to operational quality"* [19].

B. Identity Application

In addition to identity construction, the organisation is also concerned about the mechanism with which enrolled identities are accessed and used. There are four main dependent constructs that affect organisations identity access policies; **relying parties, objectives, conditions,** and **accessibility.**

a) *Relying parties*

At the most basic level, the organisation needs to specify the relying parties that require access to individuals' identity. There are two main types of relying parties: **organisational** and **individual.**

First of all, there are **intra- or inter-organisational** entities that require access to the identity. Intra-organisational access is typically a requirement since the organisation needs to create and manage identities in the first place. But access to identities within the organisation can support other functions that the organisation needs to carry out. For example, the BruNIR is not only responsible for the distribution of the identity cards in the country, but also for the monitoring of identities across borders. Recent developments have meant that the Brunei identity card can now be used as a passport at land borders with Malaysia. Therefore, the BruNIR requires other forms of internal access to support these activities.

This is not so in the Indian context, where the UIDAI was setup solely to handle the registration of identities. The main focus here lies on the inter-organisational access of identity. In its plans to introduce the identity system the UIDAI clearly established and discussed plans with several different third party organisations that include PDS, NREGS, as well as the general education and health provision systems.

In the UK, the IPS has defined both intra-organisational use of its systems (identity cards as passports), as well as its inter-organisational aims by identifying various agencies, such as law enforcement agencies and the Department of Work and Pensions. The Bruneian context, on the other hand, has comparatively ill-defined inter-organisational obligations, only stating its intent of creating a multi-purpose smart card, which can be used by any third-party organization.

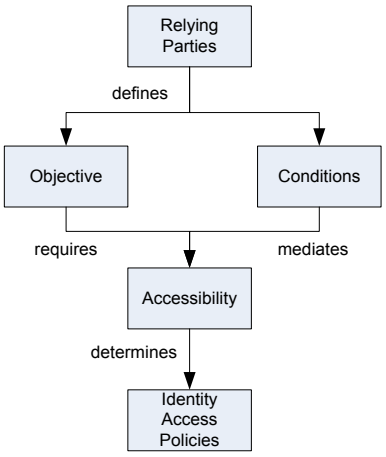


Figure 3. Organisations Identity Application requirements

In addition to organisational reliance on identity, the organisation also needs to recognise the **individual** as a relying party who may be able access his own identity and personal information. This is especially the case in the UK scenario, where IPS has specified that individuals should be able to access all their information on the system, which is envisioned to eventually be an online service [15], [20], [24]; India and Brunei have not specified any mechanisms by which individuals can directly access or view their identity records.

b) *Objectives*

Each relying party that the organisation identifies will have its own separate set of objectives. These can be expressed in terms of **enablement** and **proof**.

The use of identity to mediate the provision of services will always create a division between those who have access, and those who do not; identity systems are either primarily used to enable, or deny access. Whether the intention is to use identity to either enable or disable individuals is captured by the **enablement** construct. In India, the main intention of the relying parties is the enablement of poor citizens to access services that they have a right to, but currently do not find accessible. Additionally, Indian banks are focused on introducing new forms of mobile banking, thus enabling individuals to access new services that are to be developed.

The primary objectives in the UK context are to preventing undesirable activity (benefit fraud, crime, illegal immigration, and terrorism). The Bruneian context has described a largely enabling use of identity, with its intention to support the introduction of new on-line services introduced by third parties.

Proof describes the objective of the relying party in using individuals' identity as a single proof of identity, or as a key that enables the tracking of an individual across multiple interactions or contexts. The Indian case provides an illustration of a high-linkage scenario, where all relying parties are advised to use individuals' identity numbers as a foreign key to their own systems. It even suggests that relying parties make use of individuals' identity internally, so as to keep track of employees. The Bruneian case makes

 ISBN: 978-1-61208-176-2

no such recommendations, nor enforces any rules to such linkages, resulting in a mixed approach between parties where certain relying parties make use of the identifier as an index to their records, while others merely use the identity as a proof.

c) *Conditions*

The organisation will need to identify the conditions under which the access to the identity will take place, and may thus affect access requirements; this can be expressed as **risk level** and **timeliness.**

Risk level is a measure of the security sensitive nature of the information access. Information access that is done under high-risk conditions, such as that involving terrorism, would have greater access privileges, when compared to a low risk situation that has little implication for the country, organisation, or relying party.

The importance of risk level in the development of the identity system and the information access policies is most evident in the UK scenario. While most access by third parties would be recorded, any access for the purpose of counter-terrorism would take place without consent, and would not be recorded [15], [20], [24].

The BruNIR has created an official channel, through which law enforcement can send a written request, with supporting reasons, to obtain information. The UIDAI has not specified any direct access to the information by third parties, but, its N-IDMS plans state that one unique identifier per individual would be useful for third parties to keep track of employees that might pose a risk of corruption - for example, to track inspection officials who come into contact with food that is given out to the poor [25], or the presence of doctors and teachers ensuring they are where they need to be [14].

In addition to risk level, the **timeliness** of information access is another factor to consider. Since one of the many cited benefits of an identity system is improving efficiency, it is not surprising that the need to access information quickly is an important factor.

An example is the planned use of the UK N-IDMS for the purposes of Criminal Background Checks (CRB), which are required for persons applying for employment work in a range of sectors. Existing CRB procedures take a long time to verify individuals' identity to check their CRB status, leading to a backlog of applications. It would be beneficial if the agency carrying out these background checks could verify applicants' identity more quickly, and thus it was seen as a prime candidate for gaining some form of access to the identity system; *"the time for issuing Criminal Records Bureau disclosures could be reduced from 4 weeks to 3 days"* (ID Card Benefits Overview).

The UIDAI has also highlighted the time-sensitive nature of third parties, stressing the importance of addressing the application of current ration cards due to *"prolonged delays in processing the application"*, and the advantages in using the unique ID number in the distribution of rice grain [25]. The Brunei N-IDMS has no specific examples regarding the timeliness of information, but improved efficiency was a main factor in the introduction of the smart card system, as it would allow the transfer of information in digital format reducing the overhead for filling in forms [23].

d) *Accessibility*

Once the organisation has identified the relying parties their respective objectives, and the conditions under which they are operating, it can then go on to define the accessibility of the system to these parties. The access to the system can be described in terms of **information set, localised,** and **direction**.

Information set describes the type and amount of identity information that the relying party will have access to. For the UK system, with its emphasis on national security, certain authorities would be able to gain access to all the personal information without individuals' consent. The scenario in India is such that no relying party will have access to the personal information - the system will only confirm or deny the accuracy of personal information. . The BruNIR has stated that third-party organizations will not have any access to the database, and can only access the information that is visible on the card and stored on the smart chip.

Localised refers to the spatial mode of access to the identity system: at one end of the spectrum, a check of identity can be limited to a local point, at which an individual physically presents the identity, and at the other end is the remote access of identity through a networked database from any number of parties. The Bruneian N-IDMS does not provide third parties with any remote access to their database; all the information and authentication functions that the relying party can access, is stored on the card itself. The *raison d'etre* of the Indian N-IDMS is remote authentication, so third parties have access across a network. The UK N-IDMS specified a range of access options including local options (such as visual authentication and local chip authentication), but also fingerprint authentication across a network.

Meanwhile, the **direction** of information access describes the *push* or *pull* nature of identity access; this in turn defines the *read* (including authentication) and *write* capabilities of the relying party. The Indian N-IDMS does not provide relying parties with any ability write information to the database. The transactions are a pull of information, where the third party requests confirmation of identity. The UK N-IDMS is able to record information about the third party access when performing authentication procedures. A new entry is created on the database recording the time and location of the authentication; this represents a combined push-pull operation, where information is read from and written to the identity database.

The Bruneian N-IDMS does not provide any remote access, but law enforcement can send in a written request, which is a remote pull of information. However, third parties can store information onto the chip when required. This represents a local push of information onto the card, and therefore affects the overall information access policies that need to be set in place.

 ISBN: 978-1-61208-176-2

IV. FIT FOR PURPOSE

The previous sections have catalogued organisations' options in the construction and use of identity – and the choice of options has to match the purposes for which the system is deployed. *Who are the relying parties that require access, and what identity information does the system need to hold?* To ensure that that the system being implemented will be fit for purpose, an organisation needs to tailor the identity construction to support the requirements of those purposes.

The Indian system, with is stated purpose of enabling access to services for the poor, was quick to identify welfare agencies as relying parties, and to ensure that individuals are able to enrol (by devising the appropriate authenticity requirements for a target population that suffers from both low universality and intimacy).

In the UK, with the main purpose being the reduction of crime and terrorism, the organisation identified law enforcement agencies as a core relying party, as well as defining strict authenticity and uniqueness requirements that would support its security goals.

In Brunei, the main aims of the system were firstly to modernise their current identity infrastructure, and secondly to create a multi-function digital identity infrastructure that could be used by various third parties (especially in the provision of e-government services). As of now, there has been relatively low uptake of the system by third parties. This investigation reveals that this due to the lack of specifying relevant third parties, and thus catering for their needs and requirements. However, recent efforts to engage with a relevant stake holders in neighbouring country of Malaysia has resulted in the use of the identity cards as digital passports [26]

V. CONCLUSION AND FUTURE WORK

Using a case study approach, three different implementations of N-IDMS were examined and compared, has and this uncovered a set of choices that organisations can over **identity construction and identity use** processes. These choices must be made in line with the **purpose** of the IDMS

The organisation's requirements for identity construction will determine the amount and type of information that is collected and stored. The choice of biographical information is influenced by organisations' **authenticity** requirements, which is further mediated by the **universality** of current identity documents, and the levels of **intimacy** of the organisation to the target population. The organisation's **uniqueness** requirements influence the choice of biometric information; it is affected by the organisation's **obligations** to which it must adhere, as well as the **performance** of the biometric, which must be considered within the real-world **population** parameters.

The requirements for the use of identity will affect the identity access policies implemented. Beginning with the **relying parties** that need to access the system, the organisation must consider the various **objectives** of each party, as well as the **conditions** in which they operate. Only then will the organisation be able to specify **accessibility** of the system, and hence the identity access policies.

It should also be noted that the purpose also has an influence over the authenticity/uniqueness requirements, and vice versa. Certain purposes might require different sets of information, and the type of information within the system will place limitations on the purpose of the system; for example, a system that provides proof of age only needs to collect individuals' date of birth, whereas one designed to counter terrorism may require address information, and possibly audit trails of use.

The findings of this study further the current understanding of factors that should be considered in the design and operation of NIDMS; the codification of the identity requirements into a framework can be used to aid discussions and critiques of IDMSs. For example, [27] state that attention should be paid to issues of purpose, population scope, data scope, and users of the data. In our framework, those elements are refined into more detailed concepts and the relationships between the various concepts are elaborated. Similarly, [3] describes a short-circuiting of identity debates through the use of international obligations, language ambiguity, technological focus, and expertise. Our framework addresses these concerns by explicitly listing the considerations, thus reducing ambiguity and short-circuiting, while also introducing non-technological decisions such as relying parties, and their unique purposes.

The organisation's uniqueness consideration provides another area of comparisons to current work in the field. Drawing from [28] recommendations when implementing biometric systems, organisations should not only pay attention to the False Acceptance and False Rejection rates of biometric technology, but also consider how well they match and population characteristics, and how easy they will be to present; these are all present in the framework as sub-dimensions of the performance and environment constructs.

Therefore, the framework here serves as a guide for organisations and system designers to build effective N-IDMSs. It encourages focused debate, consideration, and definition of various critical components, ensuring that the identity information collected and the technology chosen are both fit for purpose, thus assisting in the implementation of successful identity systems.

A limitation of our current research is its emphasis on biometric identifiers; all the systems in all 3 case studies depend on them. Not all IDMS use biometrics systems, which limits the generalisability of the uniqueness framework. Future research will need to address these concerns, and further develop the framework to be applicable to non-biometric implementations.

Work will also need to be done to develop guidelines to effectively express requirements for uniqueness, authenticity and purpose; doing so will further help to increase communication in the field and encourage proper debate, while keeping the scope of the system concise and to the point.

 ISBN: 978-1-61208-176-2

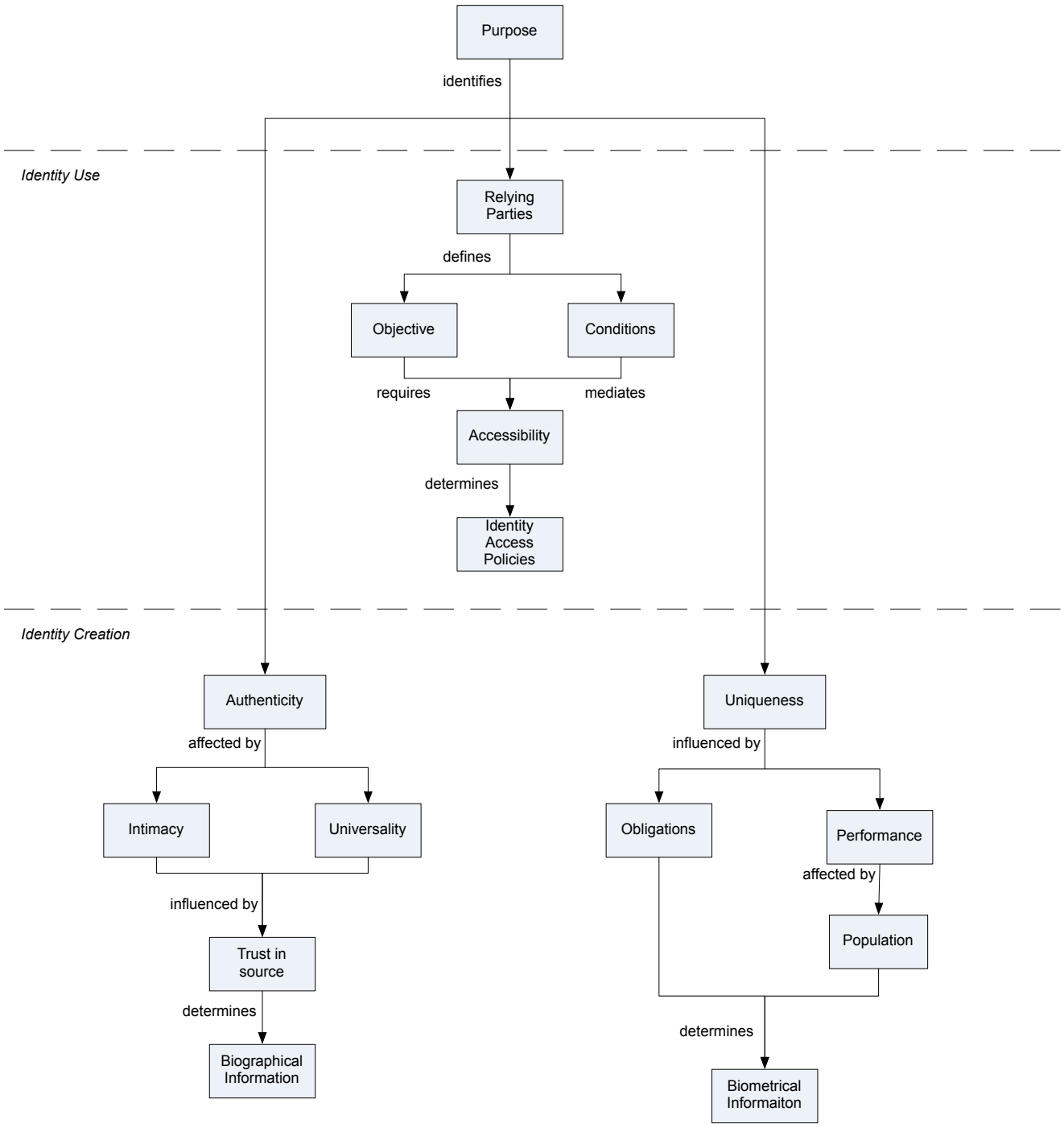


Figure 4. Complete framework displaying organisational identity requirements

 ISBN: 978-1-61208-176-2

REFERENCES

[1] A. Rahaman and M. A. Sasse, “A framework for the lived experience of identity,” Identity in the Information Society, vol. 3, no. 3, pp. 605-638, 2011.

[2] E. A. Whitley and G. Hosein, “Global identity policies and technology: do we understand the question?,” Global Policy, vol. 1, no. 2, pp. 209-215, May. 2010.

[3] E. A. Whitley and G. Hosein, Global challenges for identity policies. New York, USA: Palgrave Macmillan, 2010.

[4] BBC, “Identity cards scheme will be axed ‘within 100 days’,” BBC News, no. 2002, 2010.

[5] M. Meints, and H. Zwingelberg, "Identity Management Systems – recent developments," Deliverable 7.2, Future of Identity in the Digial Society, 2009.

[6] J. Taylor, M. Lips, and J. Organ, “Information-intensive government and the layering and sorting of citizenship,” Public Money and Management, vol. 27, no. 2, pp. 161-164, Apr. 2007.

[7] J. Taylor, M. Lips, and J. Organ, "Citizen identification, surveillance and the quest for public service improvement: themes and issues," paper to the European Consortium of Political Research ‘Privacy and Information: Modes of Regulation' Joint Session Helsinki 7-12 May 2007.

[8] H. Kubicek, “Introduction: conceptual framework and research design for a comparative analysis of national eID Management Systems in selected European countries,” Identity in the Information Society, vol. 3, no. 1, pp. 5-26, Apr. 2010.

[9] P. White, “Identity Management Architecture: a new direction,” in 8th IEEE International Conference on Computer and Information Technology, 2008, pp. 408-413.

[10] B. L. Berg, "Qualitative research methods for the social sciences," Boston: Allyn and Bacon, 2001.

[11] K. Punch, Introduction to social research: quantitative and qualitative approaches. Thousand Oaks, California, Sage Publications, 1998.

[12] J. M. Corbin and A. Strauss, Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, California, Sage Publications, 1990.

[13] Unique Identification Authority of India, UIDAI strategy overview: creating a unique identity number for every resident in India. India: UIDAI, 2010.

[14] Unique Identification Authority of India, Aadhaar handbook for registrars. India: UIDAI, 2010.

[15] Identity and Passport Service, National Identity Service: delivery update 2009. London, England: Home Office, 2009.

[16] D. Blunkett, Identity Cards: the next steps. London, England: Home Office, 2003.

[17] Identity Cards Act London, England: House of Lords, 2006.

[18] R. Yunos, “Immigration services through the ages,” Brunei Times, 01-Feb-2009.

[19] Unique Identification Authority of India, Biometric design standards for UID applications. India: UIDAI, 2009.

[20] Identity and Passport Service, Strategic action plan for the National Identity Scheme. London, England: Home Office, 2006.

[21] London School of Economics, The Identity Project: an assessment of the UK Identity Cards Bill and its implications. London, England: LSE, 2005.

[22] J. Ashbourn, Practical biometrics: from aspiration to implementation. London, England: Springer, 2004.

[23] A. Rahaman and B.I.N.R. Department, “Interview - BruNIR.” 2010.

[24] Identity and Passport Service, National Identity Scheme Delivery Plan 2008: a response to consultation. London, England: Home Office, 2008.

[25] Unique Identification Authority of India, Envisioning a role for Aadhaar in the Public Distribution System. India: UIDAI, 2010.

[26] A. Razak, “Brunei, M’sia first in SEA to use IC as passport,” Brunei Times, 2007.

[27] S. Kent and L. Millett, IDs? Not that easy: questions about nationwide Identity Systems. Washington, United States: National Academies Press, 2002.

 ISBN: 978-1-61208-176-2

TABLE I. SYSTEMS ANALYSED AS PART OF THE STUDY

	Brunei	**India**	**UK**
Population Size	**407, 000**	**1, 170, 938, 000**	**62, 218, 761**
Date Implemented	2000 – today	2010 – today	2008 – 2010 (abolished)
Purpose	Multi-function smart card	Support poor in accessing services	Prevent terrorism, crime, benefit fraud, travel card
Mandatory	18 and above	All citizens	Voluntary (mandatory for high risk personnel; airport staff, etc.)
Unique ID Number	Yes	Yes	Yes
Identity Card	Yes	No	Yes
Smart Chip	Yes	No	Yes
Centralised Database	Yes	Yes	Yes
Authentication (Against Card)	Yes	No	Yes
Authentication (Against Database)	No	Yes	Yes
Record Authentications	No	No	Yes (stored on Database)
Information Read	Third Parties can access biographical information on card and chip.	Third parties can confirm information accuracy (yes/no response).	Third parties can access biographical information on card and chip. Information can be pushed from the database to third parties. Security organisations can get access to all information on the database (through information commissioner).
Information Write	Third parties can to write to the smart card	None	Information can be pushed from third parties to the database.

 ISBN: 978-1-61208-176-2

Towards the Automatic Management of Vaccination Process in Jordan

Edward Jaser
Princess Sumaya University for Technology
King Hussein School for Information Technology
Amman, Jordan
ejaser@psut.edu.jo

Islam Ahmad
Royal Scientific Society
Information and Communication Technology
Amman, Jordan
islam@rss.jo

***Abstract*—Rural communities in developing countries are faced with many challenges due to its geographical and demographic conditions. This has been evident in many studies and surveys. Health issues are among the top priority challenges in governments' agendas. One important example is the vaccination of new born babies and young children. Vaccination is generally considered to be the most effective method for preventing infectious diseases. The rate for non-vaccination is much higher among communities in rural and remote regions. Information and Communications Technology can play an important role in assisting the government to manage the process and help reduce the rate of non-vaccination. In this paper, we describe a mobile system developed to electronically manage the vaccination process. Early evaluation demonstrates the benefits of such system in supporting government activities.**

Keywords- information and communications technology for development; health systems; mobile application; vaccination; rural areas.

I. INTRODUCTION

Governments in developing countries, and even in developed ones, face challenges with relation to services provided to communities living in rural and remote areas. Among these, health services occupy high priorities in government planning and funding. Quality health services are offered in capital and big cities. This is mainly because those cities offer more opportunities to medical staff to forward their careers in addition to the ease of life and many other advantages. This leaves rural areas and remote communities deprived of specialized and experienced medical staff. It is not difficult to imagine that many medical cases will have to travel to the capital city or other big cities to obtain needed treatment; or wait till the next medical day in their region (where a consortia of medical doctors visit rural areas) to happen. This has been a challenge even in advanced societies. A good example is [1], a study by Lenthal et al. describing the challenges facing rural Australia as a result of decreasing numbers of nurses and midwives.

Another characteristic challenge in remote areas, due to geography and the dynamic demography, is coverage. Governments face daunting task to outreach for those communities with awareness information, health warnings, medical specialists' visits and other events. In the case of Jordan, the government and related NGOs spend considerable budget to produce and print leaflets and produce TV and radio content. However, the question remains about efficiency of coverage among intended population.

Information and Communication Technologies (ICTs) are now widely considered by developing countries as the motor of growth, the driver of efficiency and effectiveness and the tool to enhance human development. With the advancement of ICTs and the Internet, communication and web-based technologies can be exploited to address many challenges with relation to improving coverage and obtaining a much needed accurate statistics and information.

In recent years there has been concrete evidence on the impact of social networking website in many aspects of life. An obvious and recent example is current events in the Middle East and North Africa. Many claim major roles for social networking tools such as Facebook and Twitter in the dynamics of these events [2]. These tools are changing the way people communicate, receive and exchange information. Such tools easily attract users as they are discrete, connect large number of individuals and eliminate the middlemen. While most popular networking websites are social in nature, professional networking websites can also be used as an efficient and cost effective tool to tackle issues and problems in society.

Many ICT interventions have been introduced to address social challenges including those of rural communities. In [3], the authors addressed the role of mobile technology and the viability of this technology in enhancing productivity, facing poverty, and improving social conditions in general. Jun [4] provided several evidences in China on the impact of mobile applications socially such as addressing employment.

One very important and priority sector is health. As mentioned earlier, quality health services are specific to large communities and adequate services or support are not available for rural areas. Health is an obvious sector that can benefit from opportunities that the technology offers as shown in many studies. In a comprehensive study [5] carried out to assess the application of ICT in health sector in terms of accessing information and disseminating awareness content in Uganda, Omona and Ikoja-Odongo concluded that there is need to support and promote ICT as the most effective tool for health information access and dissemination. The opportunities and benefits of mobile and wireless technologies for healthcare service delivery, improving patient safety and reducing cost were also the subject of a research by Ping Yu et al. [6]. The study

 ISBN: 978-1-61208-176-2

researched m-health solutions and the challenges for developing and deploying m-health applications. Maeda et al. [7] proposed a framework for mobile application for health education and awareness.

Recognizing the important role ICT can play in improving the outreach and the feedback from health services, we started a pilot project concerned with enhancing health services to women and children in remote and rural communities in Jordan. The aim of the project is to evaluate the impact of ICT on improving such services and compensate for the lack of experts and medical staff. The project contains tool for medical practitioners to interact with the public regardless of their geographical proximity. The system allows contributions from medical doctors, medical students, nurses, pharmacists and other medical personnel in Jordan to assist stakeholders (whether doctors or patients) with questions related to health issues. Also, it allows interaction between users (patients) themselves to form common interest support groups. The system's information channels, such as mobile phones allow access to health information to such groups in a cost effective manner.

In this paper, we report on one module of the project that recorded some encouraging results qualifying it to be adopted nationally. The module is concerning the management of the vaccination process of new born babies and children. The importance of such module comes from its impact on health. Most children who are not appropriately immunized are at risk of serious conditions. The automatic management of vaccination is therefore a necessary application.

II. THE CHALLENGE OF VACCINATION

Ever-changing vaccination schedules can be confusing for providers (clinics, doctors, hospitals) and parents. The Ministry of Health maintains a vaccination program that is updated and checked regularly. As soon as a new baby is born, parents are given a card with the vaccination schedule. It will be then the responsibility of the parent to follow the dates of each vaccine. Possibly, it's not an issue in urban communities with all existing electronic gadgets to remind people. Nevertheless, compounding this problem is the fact that vaccination records are often scattered. In rural areas in particular, the process is manual and records are kept on cards given to parents and on papers kept at local clinics. When records are scattered, it is difficult to assess whether a patient is up-to-date or not. It makes it harder for parents in rural and remote communities to maintain the process especially that the process is stretched over a long period of time.

On the other hand, it is also the responsibility of the clinic or the doctor to make sure that enough vaccines are stored in the local clinic to cover the need for the area they serve. Clinics usually have no statistical information on volume of vaccines needed daily or weekly. This can also lead to other challenges like storage. Clinics in rural communities are occasionally not equipped to store vaccines for long period of time.

Recent research has demonstrated specific and practical procedures medical staff can adopt to improve effectiveness in immunizing children, including the following: 1) sending parents reminders for next vaccination; 2) using printed material during calling at local clinics to remind parents and staff about importance of vaccination and vaccination table; 3) contribute to keep a statistical records on immunization rates for improvement effort.

Vaccination coverage in Jordan nationwide is relatively high [8]. Statistics show that the rate is higher in urban areas than rural.

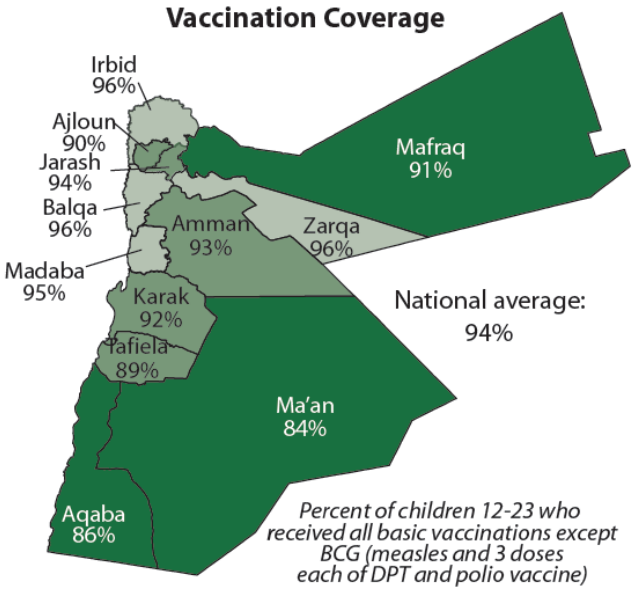


Figure 1: Vaccination coverage in Jordan (*source: Jordan 2007 Population and Family Health Survey*)

Figure 1 shows the vaccination rate among communities in rural areas (such as in the governorates of Ma'an, Tafiela and Aqaba south of Jordan) are lower than other governorates. The reason can be attributed to lack of effective awareness, illiteracy rate, lack of reminders and lack of medical personnel. Any intervention should address these issues.

III. SYSTEM REQUIREMENT

From the challenges mentioned in the previous section, an ICT intervention to manage the vaccination process will contribute to improving the vaccination rate especially among residents of rural and unprivileged areas. The high-level requirements of the system are to:

1) Register new born babies with the system and calculate the vaccines based on the vacination schedule maintained by the Ministry of Health.

2) Issue reminders to parents reminding them on the date and the type of the due vaccine.

3) Issue reminders and volume information to clinics on the number and types of vaccination they will be expecting to perform in a specific day to make sure they secure the needed quantities.

 ISBN: 978-1-61208-176-2

4) Provide awareness information to parents to help them establish the importance of the vaccines for their children.

To identify the interaction requirements of the system, we need to understand the main stakeholders and how they will use the system. We have two main stakeholders: parents and clinic staff. Following are description of the main stakeholders:

Parents: Usually parents in rural communities are not exposed to technology such as internet and all the tools that come with it. In a survey conducted prior to the design of the ICT intervention to investigate the best way for the system to interact with those users, it was noticed that more than 90% of the surveyed users own at least one mobile phone. This is quite significant penetration rate. Most of these phones are basic ones. The usage of mobile phones is for the purpose of making and receiving calls as well as communication through text messages. Our conclusion was that any project should have a mobile component to communicate information with the users.

Clinics: When we mentioned clinics servings rural and remote communities, we are assuming basic infrastructure. No internet or computer. Some of the clinics are even mobile clinics to provide services to the moving population (Bedouins) and they tend to have minimum equipments. Any solution to be adapted nationwide should take into consideration the cost factor i.e. minimum is to be spent on infrastructure. Clinic staff should have the option to interact with the system using the internet if possible, or using smart mobile devices which are cheap to acquire and install.

IV. MOBILE APPLICATION

Given the identified requirements, we designed the vaccination management system based on a clear identified scenario and simple workflow.

Workflow: Clinic staff registers new born babies with the system. This can be done either using the internet (website) or, if the infrastructure is not there, using smart phone over 3G networks. The application can be downloaded and used with any java enabled phone. The clinic personnel need to capture basic information about the child (name, date of birth, weight, height and contact details of parents) and send the information to the system as an SMS message. Once the information of a new child have been received and stored, the system uses the vaccination schedule issued by the Ministry of health to calculate the dates of the vaccines for the registered child and store them on the database. The system continuously checks the database to produce a report of which children due to be vaccinated in a certain day. The system automatically sends the parents a reminder on the next vaccine for their child and at which clinic. Also, the system sends statistical and volume information to clinics about expected children and vaccines at a certain date. When the vaccination of a child takes place, the clinic personnel send this information to the system to maintain the child record. Figure 2 summarizes the workflow of the system.

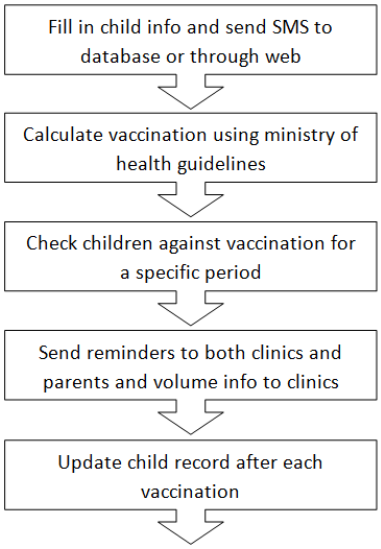


Figure 2: Work flow of the system.

The database maintains up-to-date record of vaccination to be used by decision makers at the ministry to obtain information for the purpose of reporting and planning.

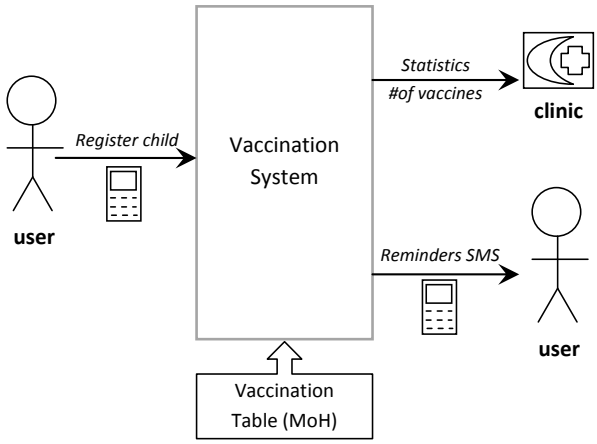


Figure 3: Architectural design for the Automatic vaccination system

Figure 3 depicts the architectural design of the system, the input/output channels and support information. Figure 4 provides more details about the process and the modules in the system.

Objectives: The main aim behind the design of the pilot system has been to measure the advantages and impact of ICT interventions in enhancing vaccination process and support health clinics and hospitals in rural and remote communities. During the life of the project we attempted to answer the following key research questions: (1) Could ICTs contribute to the enhancement of the general health of rural and remote communities? and (2) What is the minimum infrastructure needed for the deployment of the automatic vaccination management system?

 ISBN: 978-1-61208-176-2

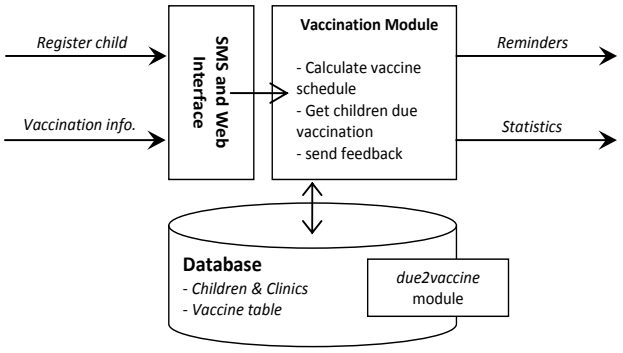


Figure 4: Main components of vaccination system

Implementation: Our intervention was developed using the following technologies: Java Server Pages (JSP) and Java Servelet for the Dynamic Interaction with the users in the user interface; J2ME to develop mobile application; Hyper Text Mark-up Language (HTML) for the Content of the static pages in the system; JavaScript for the validation of the data entered by the users; AJAX for the smooth interaction in items (Data Swap without DB Rendering); The Tag-libs technology for the Modularity and Template adaptation in Inner and Main Pages; MySQL for database functionality; and XML for the optimized and structured data transfer.

V. EVALUATION

We made the necessary awareness about the availability of our system to concerned stakeholders and we started to collect the needed information to answer our research questions. Five clinics in Jordan were selected serving communities in remote and rural areas. The selection was made together with the Ministry of Health based on need analysis. We provided these clinics with minimum infrastructure required (i.e. a netbook and a smart mobile phone per clinic). Valuable information were collected (either by interviews and/or questionnaires) from various stakeholders with this regards. The evaluation at this early stage is subjective and is based on user experience with the system. We noticed that all feedback from clinics were positive. The system allowed them to plan for both number of staff needed and number of vaccines required. Several comments were made about the use of mobile devices as a data entry device compared to using a laptop in terms of easiness and correctness. Parents who tried the system expressed their satisfaction on receiving reminders about the process. Parents are only receiving a couple of line text in the form of SMS and don't need to report or perform any data entry. Basic mobile phones were suitable for the task.

The subjective evaluation showed that user experience from using the system was smooth and realized with satisfaction. However, and in order to obtain a clearer picture in terms of benefits on how this module help tackling and important issue which is children missing vaccination (higher in rural areas), a wider deployment is needed nationwide once a decision is made to adopt the solution.

Overall, and in spite of the advancement in mobile applications developments and innovations, there are certain challenges: (i) Concrete evidence: we are in need of a robust analysis and evaluation tools and standards of mobile intervention in health to help designing better and effective services; (ii) Legislations: this is quite important to establish a clear policies and laws to govern ICT interventions and their deployment; (iii) Sustainability of ICT interventions, sustainability is important issue for ICT4D projects. There should be clear understanding of how to fund these projects from both public and private sectors and to continue to provide resources and (iv) Capacity building: there should be focus on building the competency of various stakeholders in terms of ICT usage. Also mobilize resources to bridge the technological gap facing rural communities.

VI. CONCLUSION AND FUTURE WORK

Deploying the pilot vaccination management system in selected clinics serving remote communities demonstrated the impact ICT interventions can have on enhancing services and its outreach. More research and investigation is needed to deal with smooth and effective online communication between patient and health workers. Future work will focus on widening the evaluation to include more clinics and region in rural Jordan to reach a working system that can be adopted nationally.

REFERENCES

[1] Sue Lenthall, John Wakerman, Tess Opie, Sandra Dunn, Martha MacLeod, Maureen Dollard, Greg Rickard, and Sabina Knight; "Nursing workforce in very remote Australia, characteristics and key issues". Australian Journal of Rural Health, 19: pp. 32–37. 2011.

[2] Marko Papic and Sean Noonan; "Social Media as a Tool for Protest" web article http://www.cfr.org/democracy-and-human-rights/stratfor-social-media-tool-protest/p23994, visited on the 20th January 2012.

[3] Lisa Cespedes and Franz Martin; "Mobile Telephony in Rural Areas: The Latin American perspective"; The i4d print magazine; Vol. VII No. 9 pp. 10-11, January-March 2011.

[4] Liu Jun; "Mobile Social Network in a Cultural Context". In E. Canessa and M. Zennaro: m-Science: Sensing, Computing and Dissemination. ISBN:92-95003-43-8. Pp. 211-240. 2010.

[5] Walter Omona and Robert Ikoja-Odongo; "Application of information and communication technology (ICT) in health" Journal of Librarianship and Information Science; 38:pp. 45-55. 2006.

[6] Ping Yu, Mingxuan Wu, Hui Yu, and GC Xiao; "The Challenges for the Adoption of M-Health" IEEE International Conference on Service Operations and Logistics, and Informatics, 2006. pp. 181-186, 21-23 June 2006.

[7] Toshiyuki Maeda, Tadayuki Okamoto, Yae Fukushige, and Takayuki Asada; "Mobile Application Framework for Health Care Education" 7th IEEE Consumer Communications and Networking Conference, pp. 1-2. January 2010.

[8] Department of Statistics; "Jordan 2007 Population and Family Health Survey: Key Findings Department of Stattics". Amman – Jordan; 2007.

 ISBN: 978-1-61208-176-2

Three Dimensional Printing:

An introduction for information professionals

Julie Marcoux
Bibliothèque Champlain
Université de Moncton
Moncton, Canada
Julie.Marcoux@umoncton.ca

Kenneth-Roy Bonin
School of Information Studies
University of Ottawa
Ottawa, Canada
Kenneth-Roy.Bonin@uottawa.ca

***Abstract* - Advanced by some as the next great emerging technology to enjoy overwhelming market penetration, three dimensional (3D) printing could have significant information implications, notwithstanding limited coverage in the information science literature. This review of complementary material from other sources provides the introductory definitions, technical descriptions and indications of future developments relevant to information professionals.**

Keywords - 3D printing; three dimensional printing; additive fabrication; digital fabrication; rapid prototyping.

I. INTRODUCTION

Three dimensional (3D) printing has the potential to impact the transmission of information in ways similar to the influence of such earlier technologies as photocopying and telefacsimile. This review identifies sources of information on 3D printing, its technology, required software and applications. Although the subject initially may seem to be of particular interest to engineers, efforts have been made to identify resources relevant to exploring the implications of 3D printing technologies for those working in the information sciences: librarians, archivists, museum collection specialists, and managers of documentation centers and information services in the public and private sectors. Accordingly, the following presentation provides definitions, reports the results of a literature review, explains the technology, and outlines directions for future work.

II. 3D PRINTING DEFINED

All sources identified through a literature search on the subject of 3D printing shared the common characteristic of providing a definition. The fundamental idea varies little from one source to another. Most agree that 3D printing consists of downloading a blueprint or a special computer file to a printer capable of 'printing' sophisticated three-dimensional objects through an additive process that 'prints' layers of material [1]-[10].

Bradshaw et al. [3] distinguish three fundamental methods for fabricating objects: 1. cutting the object out of a block of material; 2. creating a mold and then filling it to create the object; or 3. adding shapes together in order to make an object. The technology of 3D printing falls into the latter category, in which objects with moving parts can be created, impossible using the two other methods alone [3] [9] [11]. There is however, considerable diversity in actual 3D printing production.

The different techniques have led to certain semantic disputes. Wiegler [10] quotes a few researchers who feel that the term should be reserved for the particular 3D printing technique created by Zcorp, the company credited with creating a cheaper 3D printing technique in which a nozzle similar to a 'glue gun' is used to print out objects. Others believe that the term '3D printing' should be used generically, to include all types of additive manufacturing, because it is easily understandable by the general public [10]. It is this latter connotation which will be adopted in this review, employing the term '3D printing' to encompass all techniques that lead to a three dimensional object being printed, including such variations as 'selective laser sintering' and 'fused deposition modeling', which will be explained later.

II. LITERATURE REVIEW

As summarized by Weinberg, “the line between a physical object and a digital description of a physical object may (...) begin to blur. With a 3D printer, having the bits is almost as good as having the atoms” [9]. Librarians and information professionals, well aware of the important contributions of other electronic technologies to the disintermediation of information, should be predisposed to understand the implications of 3D printing. To assess the prevalence of articles on 3D printing in the information science literature, a three phase search for relevant articles was conducted. The first phase consulted three information science databases: *Library and Information Science Abstracts*; *Library, Information Science and Technology Abstracts*; and *Library Literature & Information Science Full Text*. The search terms employed were '3D print*', 'three-dimens* print*', 'three dimens* print*' and 'tridimens*print*'. When search results proved disappointing, a more general literature search was conducted in a second phase of the literature review. This involved a greater variety of sources, including reports and conference proceedings, newspapers, industry publications,

 ISBN: 978-1-61208-176-2

electronic media and online information, and articles in engineering databases. When identified, relevant references were subsequently accessed to obtain a more comprehensive picture of the subject. Supplementary synonyms for the term '3D printing' found in this broader literature were carefully noted so that the initial three information science databases could be interrogated again in a third phase of the literature review.

Results of the three literature searches reveal that most of the relevant material on 3D printing has been published within the last six years, including many sources less than two years old. As a technology, however, 3D printing has been around for some time, and commercial printers "have existed for years" [10] [12] [13]. Bradshaw et al. [3] confirm that the first patent was deposited in 1977. One reason for the recent nature of most of the literature is that prices for 3D printers have dropped sufficiently that individuals can now afford to purchase their own equipment [1] [4] [6] [13]. This has encouraged greater interest in the possibilities of the technology.

The various sources of information frequently approach 3D printing as a subject quite differently. Newspaper articles provide general overviews of the subject, the 'meta subjects' and related topics. While they describe printing techniques, they rarely employ scientific terms such as 'fused deposition modeling', for example. This can lead to confusion when the full range of 3D printing possibilities is not described. Reports provide more in-depth and accurate descriptions. Blogs generally post the most recent developments, but commercial blogs unsurprisingly tend to concentrate on what specific 3D printing companies have to offer. Manufacturing company websites tend to provide a considerable amount of information about 3D printing as a process, including detailed technical information. Those that print from files created by customers even provide links to web pages for non-commercial software. Unfortunately, meta subjects and related topics are rarely covered. Engineering articles range from a tight focus on elements of 3D printing as a process, such as improving the viscosity of material to be printed, to more general considerations of meta subjects. The *Rapid Prototyping Journal* is a particularly valuable source of engineering articles in this regard.

Given the information implications of the resources identified through the more general literature search, the initial search of information science databases was repeated to take into account the terms identified in the more general literature search. This search yielded only four additional results. Two of these articles [14] [15] were by the same author, giving a brief assessment of one commercial 3D printer and of a particular piece of modeling software. Another [16] discussed combining two databases to obtain a 3D printable file of the outline of buildings in Norway. The three articles were very focused on their specific topics, and none discussed the information implications of 3D printing.

Although also narrowly focused, the fourth article [17] does describe an application of 3D printing with information implications. It discusses the use of 3D scanners and a 3D printer to create replicas of wooden stamps. The article concludes by explaining that the stamps are now easier to share with other libraries and museums, an illustration of a potential contribution 3D printing might make to extending access while preserving original archival and museum materials.

Articles relevant to the information implications of 3D printing were also discovered in the more general search conducted in the second phase of the literature review. A kinematic library was identified which has made 3D printable files of kinetic models available online [18]. The metadata scheme developed for this library might have served as a good starting point for reflecting on the classification and cataloguing of 3D printable files, but unfortunately, it does not appear to be systematically maintained, and many of the supplied links are broken [19].

Two engineering articles also found in the same general literature search outlined attempts to create classification schemes for 3D printing. Ingole et al. [20] make relevant observations about the need for more formalized standards for 3D printing and discuss some of the difficulties associated with the proprietary standards associated with commercial machines. Certain components of the proposed classification scheme, such as the use of single digits to code subjects, would have benefited from input by information professionals. The second engineering article, by Mortara et al. [21] demonstrates an awareness of such important classification concepts as faceted classification, but the proposed classification scheme is clearly aimed at engineers, and would not be easy to use for publicly accessible repositories.

III. TECHNICAL OVERVIEW

This section explains technical aspects of 3D printing found in reviewed sources. The utility of 3D printing rests in its capacity to cheaply print complex objects, such as already linked up chain-mail, using a variety of materials [2]. Certain accounts concentrate exclusively on a single 3D printing application. Seulin's article [17] on wooden stamps is an example. While these resources explain the process involved in creating a 3D object, they often focus on a specific printing technique and on a specific 3D printer.

Areas of interest which have used 3D printing to create objects include aeronautics, architecture, automotive industries, art, dentistry, fashion, food, jewelry, medicine, pharmaceuticals, robotics and toys [2] [12] [22]. 3D printable files of physical models of educational concepts would also interest academic and school libraries. Knapp et al. [23] explain that commercial models "are expensive", and give the example of "an anatomical model of the heart" which can "cost up to $600." A 3D printer can be purchased for under $1,000 and materials for "even the largest model (...) would likely cost under $50" [23]. As noted previously,

 ISBN: 978-1-61208-176-2

the preservation of artifacts is another potential use of 3D printing that would be of interest to information professionals [7] [23].

Both open-source and proprietary software may be used to acquire digital files of 3D objects during the data acquisition phase of production. Lipson and Kurman [6] credit "the emergence of cheaper, and increasingly accessible computer aided design software (CAD)" for the increasing interest in 3D printing. A number of authors mention the use of 3D scanning, which uses basic cameras and freely provided software, rather than commercial systems, to obtain digital data of existing objects [2] [9] [17] [22].

Authors also confirm that there is no agreement on file formats for 3D printing [6] [20]. The range in the types of files that currently exist include PLY files, ObjDF files, RP files, STL files VRML files and ZPR files [4] [14] [23]. Inacu et al. [23] believe that STL "is, and for the foreseeable future, will be the standard mode of data exchange in the Rapid Prototyping industry." Software used to create printable 3D files includes modeling software, file converters, model 'repair' software to clean-up files, and path generating software.

Help is also available online for anyone interested in creating printable 3D files. Google Sketchup [24], for example, does not require a user fee to access a basic level of service. Turning a model into a printable 3D file involves following instructions from a free tutorial provided by a for profit company called Shapeways [25].

Both Cornell University and the University of Bath have designed open-source 3D printers which are widely recognized for making all 3D printers more affordable: Fab@home and RepRap [7] [13]. To acquire one of these open-source 3D printers, interested parties obtain the basic building materials, follow construction instructions shared on Wikis, and then purchase printing supplies [7]. Bath even allows the commercial resale of its printers. Lipson and Kurman [6] note that Creative Commons initiatives have been inspired to work on open-source hardware licenses. Fab@home's ultimate objective is to build a machine capable of producing "complete, integrated, functional electromechanical devices" [7]. The goal for RepRap is to enable it to replicate itself by printing out all parts for a new RepRap [6].

The cost of a RepRap printer in 2010 was about $525, and it could replicate 50% of itself [13]. The operating cost of 3D printing materials can be less than $1 per cubic inch [23]. Caution must be exercised when quoting published costing information, however, since prices have been dropping so fast that they are quickly outdated. Wiegler [10] cites the example of a professor who bought a commercial 3D printer in 2005 for $31,000, noticed that the price had dropped to $19,000 in 2008, and speculated that it would drop to $10,000 by 2013.

Commercial Printers that use more advanced techniques to print objects are usually equipped with proprietary software [14]. Companies that sell 3D printers include 3D Systems, Objet Geometries, Solido LTD, Stratasys and Z Corp [2] [18]. Lipson and Kurman (2010) report that both Hewlett Packard and Xerox "are investing in 3D printing research and technology development" [6].

Several types of material can be used to print objects. Various printers handle a variety of materials, and some can even produce objects using more than one type of material. While the most commonly used materials are plastic, metals and ceramics, more exotic materials, such as chocolate, may also be used [6] [7].

There are two major variations in printing techniques. In one instance, material is deposited on a surface, and the depositing implement of the printer pulls up after a layer has been deposited in order to deposit the next layer [2]. Support material might be put in place to protect the structural integrity of the object, but must be removed later [15]. In another instance, a layer of powder or liquid is present on the printing surface. A binding technique is used to change parts of the powder or liquid in the first layer of the object. The printing surface is lowered, more powder or liquid is added, and the process is repeated to form the next layer [26]. The remaining material supports the weight of the object as it is built. Binding techniques include adding glue to material, adding an 'ink' that solidifies when exposed to ultra-violet light, or using a laser to bind material [2] [6].

Although different techniques have specific names, a semantic shift in the terms used to describe generic 3D printing has resulted in a number of variations. Selective laser sintering, fused deposition modeling and stereolithography are among the most often mentioned techniques [4] [20] [26] 27]. Stereolithography uses a liquid polymer bonded by a laser; laser sintering uses a powder which is also bound by a laser; while fused deposition modeling simply deposits material on a printing surface [28]-[30]. Certain techniques also require post-processing of the printed objects in order to solidify them or to improve their appearance [26] [31]. Post-processing steps can include 'bed manipulation', which entails forcing a change to all the material that will only have an effect on bonded parts, removing powder or support materials, heating the object, or dipping it into something else (infiltration) [26].

IV. CONCLUSION AND FUTURE WORK

Not all technical information about 3D printing could be shared in this introduction of the subject. Documenting the technology, very much a work-in-progress, must also recognize that not all authors agree on the likelihood of 3D printing gaining wider dissemination into the homes of individuals [10]. Also, as a still emerging technology, 3D printing is not without its problems, such as slow printing speeds [6]. Nevertheless, as prices are decreasing, the number of 3D printers sold worldwide has been growing steadily. And as market penetration increases, the information implications of 3D printing technologies will

 ISBN: 978-1-61208-176-2

expand as well. These include legal considerations and parallels associated with the spread of desktop computers [6] [7] [9]. Published works related to the information economy [32] [33], the democratization of manufacturing [34], and on the concept of the 'long tail' [2] will also assume greater significance [6].

The lesson learned from this initial effort to introduce 3D printing to information professionals is that explanations of the technology will not, as yet, be found in their professional literature. Hopefully, however, as they begin to appreciate the potential of desktop 3D printing technology, information professionals will have more to contribute to a greater understanding of its implications.

REFERENCES

[1] American Public Media, "Brave new world of 3D printing," [Podcast], Marketplace Tech Report, November 29, 2010. Retrieved from http://marketplace.publicradio.org/display/web/2010/11/24/tech-report-the-brave-new-world-of-3d-printing/ 25.11.2011

[2] A. Anderson. "A Whole New Dimension: Rich Homes Can Afford 3D Printers," The Economist, November 15, 2007. Retrieved from http://www.economist.com/node/10105016?story_id=10105016 25.11.2011

[3] S. Bradshaw, A. Bower, and P. Haufe. "The Intellectual Property Implications of Low-cost 3D printing," *SCRIPTeD*, vol. 7, (1), 2010, pp. 5-31, doi: 10.2966/scrip.070110.5.

[4] C. Inacu, D. Inancu, and A.Stanciou, "From CAD Model to 3D Print Via 'STL' File Format," Fiabilite si Durabilite = Fiability & Durability, vol.1, 2010, pp. 73-80.

[5] G. Lacey, "3d Printing Brings Designs to Life," Technology Education, 2010, pp. 17-19.

[6] H. Lipson and M. Kurman, Factory at Home: The Emerging Economy of Personal Manufacturing. Washington: U.S. Office of Science and Technology, 2010, n.p.

[7] E. Malone and H. Lipson, "The Factory in Your Kitchen," 2007 World Conference on Mass Customization & Personalization (MCPC), Cambridge, MA: MCPC 2007. Retrieved from http://ccsl.mae.cornell.edu/papers/MCPC07_ Malone.pdf 24.11.2011

[8] D. Smock, "Lower Prices Drive 3-D Printer Market", Design News, May, 2010, n.p.

[9] M. Weinberg, "It Will Be Awesome if They Don't Screw It Up: 3D Printing, Intellectual Property, and the Fight Over the Next Great Disruptive Technology," Public Knowledge, November, 2010, pp. 1-15. Retrieved from http://www.publicknowledge.org/files/docs/3DPrintingPaperPublicKnowledge.pdf 24.11.2011

[10] L. Wiegler, "Jumping Off the Page," Engineering & Technology, vol. 3, 2008, no. (1), pp. 24-26.

[11] C. Major and A. Vance, Desktop manufacturing [Video file]. The New York Times, 2010. Retrieved from http://video.nytimes.com/video/2010/09/13/technology/1248068999175/desktopmanufacturing.html 24.11.2011

[12] D. L. Bourrel, M. C. Leu, and D. W. Rosen, Roadmap for Additive Manufacturing: Identifying the Future of Freeform Processing. Austin, TX: University of Texas, Laboratory for Freeform Fabrication, 2010.

[13] G. Stemp-Morlock, "Personal Fabrication: Open Source 3D Printers Could Herald the Start of a New Industrial Revolution," Communications of the ACM, vol. 53, 2010, no. 10, pp. 14-15, doi:10.1145/1831407.1831414.

[14] S. Ellerin, "The Art and Science of 3D Printing," Emedia, vol. 17, 2004, no. (5), pp. 14-15.

[15] S. Ellerin,"Strata 3D Pro," EMedia, vol. 17, 2004, no. 6, pp. 28-29.

[16] J. O. Nygaard, "Semiautomatic Reconstruction of 3D Buildings from Map Data and Aerial Laser Scans," Journal of Digital Information Management, vol. 2, 2004, no.4, pp. 164-170.

[17] R. Seulin, C. Stolz, D. Fofi, G. Million, and F. Nicolier, " Three Dimensional Tools for Analysis and Conservation of Ancient Wooden Stamps," Imaging Science Journal, vol. 54, 2006, pp. 111-121, doi: 10.1179/174313106X98755.I.

[18] K. Walker and J. M.Saylor, "Kinematic Models for Design Digital Library," D-Lib Magazine, vol. 11, 2005, no. 7. Retrieved from http://www.dlib.org/dlib/july05/07featuredcollection.html 24.11.2011

[19] CTS Metadata Services, KMODDL application profile, 2004. Retrieved from http://kmoddl.library.cornell.edu/aboutmeta2.php 24.11.2011

[20] D. Ingole, A. Kuthe, T. Deshmukh, and S. Bansod, "Coding System for Rapid Prototyping Industry," Rapid Prototyping Journal, vol. 14, 2008, no. (4), pp. 221-233.

[21] L. Mortara, J. Hughes, P. S. Ramsundar, F. Livesey, and D. R. Probert, "Proposed Classification Scheme for Direct Writing Technologies. Rapid Prototyping Journal, vol. 15, 2009, no. 4, doi:10.1108/13552540910979811.

[22] S. Summit, Ok, so you can create anything. Now what? [Video file] Singularity University: Preparing Humanity for Accelerating Technological Change, 2010. Retrieved from http://www.youtube.com/watch?v=6lJ8vId4HF8 24.11.2011

[23] M. E. Knapp, R. Wolff, and H. Lipson, Developing printable content: A repository for printable teaching models, n.d.. Retrieved from http://www.3dprintables.org/printables/images/d/d7/3Dprintables_paper_final.pdf 24.11.2011

[24] Google Sketchup, 3D modelling for everyone, n.d. Retrieved from http://sketchup.google.com/ 24.11.2011

[25] Jed, SketchUp STL export tutorial, In Shapeways, n.d. Retrieved from http://www.shapeways.com/tutorials/sketchup_3d_printing_export_to_stl_tutorial 24.11.2011

[26] B. R. Utela, D. Sorti, R. L. Anderson, and M. Ganter, "Development Process for Custom Three Dimensional Printing (3DP) Material Systems," Journal of Manufacturing Science and Engineering, vol. 132, pp. 1-9, doi: 10.1115/1.4000713.

[27] I. Serban, I. Rosca, and C. Druga, "A Method for Manufacturing Skeleton Models Using 3D Scanning Combined with 3D Printing," In Annals of DAAAM for 2009 and Proceedings of the 20th International DAAAM Symposium, vol. 20, no. 1, Vienna, Austria: DAAAM International, 2009, pp. 1319-1320.

[28] Materialise, About fused deposition modeling, n.d. Retrieved from http://www.materialise.com/fused-deposition-modelling 24.11.2011

[29] Materialise, About our laser sintering prototyping service. Retrieved from http://www.materialise.com/laser-sintering-prototyping 24.11.2011

[30] Materialise, About our stereolithography prototyping service. Retrieved from http://www.materialise.com/Stereolithography 24.11.2011

[31] T. Ringdahl, "3d Printer Lets Designers Run with Shoe Design," Machine Design.com, March 19, 2009, pp. 58-59.

 ISBN: 978-1-61208-176-2

[32] A. Fenner, “Placing Value on Information,” Library Philosophy and Practice, vol. 4, 2002, no. 2, pp. 1-6.

[33] C. M. Gayton, “Legal Issues for the Knowledge Economy in the Twenty-First Century,” Vine, vol. 36, 2006, no. 1, pp. 17-26.

[34] E. V. Hippel, Democratizing Innovation, Cambridge, MA: MIT Press, 2005.

 ISBN: 978-1-61208-176-2

Unsupervised Personality Recognition for Social Network Sites

Fabio Celli
CLIC-CIMeC
University of Trento
Italy
fabio.celli@unitn.it

***Abstract*—In this paper, we present a system for personality recognition that exploits linguistic cues and does not require supervision for evaluation. We run the system on a dataset sampled from a popular Social Network: FriendFeed. We adopted the five classes from the standard model known in psychology as the "Big Five": extraversion, emotional stability, agreeableness, conscientiousness and openness to experience. Making use of the linguistic features associated with those classes the system generates one personality model for each user. The system then evaluates the models by comparing all the posts of one single user (users that have only one post are discarded). As evaluation measures the system provides accuracy (measure of the reliability of the personality model) and validity (measure of the variability of writing style of a user). The analysis of a sample of 748 Italian users of FriendFeed showed that that the most frequent personality type is represented by the model of an extravert, insecure, agreeable, organized and unimaginative person.**

***Keywords-Social Network Sites; Personality Recognition; Information Extraction; Natural Language Processing*.**

I. Introduction and Related Work

Personality is a crucial aspect of social interaction. Under the computational perspective it can be very useful for marketing and for interesting tasks such as stylometry and sentiment analysis. Recent studies showed that there is a connection between the personality of individual users and their behavior online (see Amichai-Hamburger and Vinitzky [1]). Social Network Sites (SNSs henceforth, see Boyd and Ellison [2] for definitions and history) are huge, virtually infinite, corpora where authors (users) and sentences (posts) are found together. Many scholars used data from social networks for personality classification. In 2006 a pioneering work by Oberlander et al. classified four traits of blog authors' personality using n-grams as features. Some very recent works such as Quercia et al. [10] and Golbeck et al. [4] predicted personality of users from social network data. In particular Golbeck et al. predicted personality from some users' profiles on Facebook using machine learning techniques. Golbeck's work is supervised because it required that subjects completed a personality test for evaluation. Here we introduce a novel technique for personality recognition that does not require subjects.

In the following section, we will present a system that builds on the fly one personality model for each user in a corpus in an unsupervised way and performs automatic evaluation of the models comparing all of his/her posts. Then, in Section 3, we will present the results of the analysis of personality on FriendFeed. In Section 4, we will conclude introducing possible directions for future works.

II. Unsupervised Personality Recognition

The large amount of data available from Social Network Sites allows us to predict users' personality from text in a computational way, but there are at least four nontrivial problems:

1) The definition of personality, which is a very fuzzy and subjective notion;
2) The annotation of personality in the data from SNSs, that would require personality judgements by the author themselves or by other native speakers.
3) The construction of one model for each user in the dataset.
4) The evaluation of personality models.

In the next paragraphs we are going to discuss the solutions for those problems we adopted for the unsupervised personality recognition system.

A. Definition of Personality

Psychologists describe personality along five dimensions known as the "Big Five" (see Goldberg [5]), a model introduced by Norman in 1963 [8], obtained from factor analysis of personality description questionnaires that has become a standard over the years. The five dimentions are the following:

- Extraversion (E) (sociable vs shy)
- Emotional stability (S) (calm vs insecure)
- Agreeableness (A) (friendly vs uncooperative)
- Conscientiousness (C) (organized vs careless)
- Openness (O) (insightful vs unimaginative)

Those dimentions can be represented computationally as continuous numerical variables with 2 poles: one positive

 ISBN: 978-1-61208-176-2

(1) and one negative (0). Once we have the numerical values for each attribute (one attribute is one dimention in the "Big Five"), we can easily calculate whether a user has one trait of personality (y) or not (n) or we have no information about that trait (o). From this representation, we can formalize a personality model for each user simply taking the majority class for each attribute/dimension from all posts the user made. In the end personality models are formalized as string of five charachters: one for each attribute, which one can take three possible vaues: positive (y), negative (n) or balanced (o). For example a the string `ynooy` is the model of an extravert, nervous and open-minded user.

B. Dataset

The dataset is a sample of 748 Italian FriendFeed users (1065 posts). It is a subset of the dataset sampled by Celli et al. [3]. The dataset has been collected from FriendFeed public URL, where new posts are publicly available. The dataset was already processed with a language identifier, whose performance is correct at 88%. This made easier the extraction of the Italian subset.

Our unsupervised system does not require direct annotation of the dataset, but just a set of correlations between linguistic factors and personality traits to build models. Either Mairesse et al., Golbek et al. and Quercia et al. report sets of correlations between some cues and the dimentions of personality in the "Big Five". In our system we used a set taken from Mairesse et al. because it is the largest one and it is more focused on linguistics.

C. Building the Personality Model

Mairesse et al. provides a long list of correlation coefficients between linguistic factors and the personality traits. These coefficients are obtained from an essay corpus where authors and external observers provided personality judgements following the "Big Five" model. In order to develop an unsupervised personality recognition system we need to turn those coefficients into features that can be automatically extracted from text. Among those linguistic factors that correlates with certain aspects of personality there are some regarding topic (for example if a person writes about job, leisure, music, other people), some regarding word usage (for example the frequency of words used, the use of negative particles, first person pronouns, fillers, swares) and some regarding psychological aspects (for example age of acquisition of the word used, length of the words used, expression of positive and negative feelings). Factors are supposed to be valid for the western culture. We picked up and adapted 22 features from Mairesse et al. They are:

1) **all punctuation** (ap): the count of `. , ; :` in the post,
2) **commas** (cm): the count of `,` in the post,
3) **reference to other users** (du): the count of the pattern `@` in the post,
4) **exclamation marks** (em): the count of `!` in the post,
5) **external links** (el): the count of external links in the post,
6) **first person singular pronouns** (im): the number of first person singular pronouns in the post,
7) **negative particles** (np): the count of negative particles in the post,
8) **negative emotions** (ne): the count of emoticons expressing negative feelings in the post,
9) **numbers** (nb): the count of numbers in the post,
10) **parenthesis** (pa): the count of parenthetical phrases in the post,
11) **positive emotions** (pe): the count of emoticons expressing positive feelings in the post,
12) **prepositions** (pp): the count of prepositions in the post,
13) **pronouns** (pr): the count of pronouns in the post,
14) **question marks** (qm): the count of `?` in the post,
15) **long words** (sl): the count of words longer than 6 letters in the post,
16) **self reference** (sr): the count of first person (singular and plural) pronouns in the post,
17) **swears** (sw): total count of vulgar expressions in the post,
18) **type/token ratio** (tt): defined in the formula below,
19) **word count** (wc): words in the post,
20) **first person plural pronouns** (we): count of first person plural pronouns in the post,
21) **second person singular pronouns** (yu): count of second person singular pronouns in the post,
22) **mean word frequency** (mf): simple mean of the frequency of words in the post, defined in the formula below.

$$tt = \frac{w - T}{T} \quad mf = \frac{\Sigma wf}{T}$$

where w is the count of words already used in the sentence, T is the total word count in the sentence and wf is the frequency count of the word in the dataset. Table I (from Mairesse et al.) shows how the linguistic features used correlate with personality traits. First the system extracts a random sample of the dataset for statistical purposes. The size of the sample can be decided a-priori, in this case we sampled 500 posts. From this sample the system extracts mean and standard deviation for each feature. The mean word frequency (feature mf) in this case is calculated using an external corpus of Italian (CORISsmall, see [11]) but in principle it can be calculated also from the dataset itself as relative frequency. Results are summarized in Table II. In the second step the system processes the entire dataset building a personality model for each post applying the following rules: if a sentence shows a feature correlating positively with one personality trait and the frequency of that feature is higher than mean plus standard deviation for that feature,

 ISBN: 978-1-61208-176-2

F.	E	S	A	C	O
ap	-.08**	-.04	-.01	-.04	-10**
cm	-.02	.01	-.02	-.01	.10**
du	-.07**	.02	.01	.01	.06**
el	-.05*	-.02	-.01	-.03	.09**
em	-.00	-.05*	.06**	.00	-.03
in	-.04*	.01	-.01	-.03	-.01
im	.05*	-.15**	.05*	.04	-.14**
np	-.08**	.12**	.11**	-.07**	.01
ne	-.03	-.18**	-.11**	-.11**	.04
nb	-.03	.05*	-.03	-.02	-.06**
pa	-.06**	.03	-.04*	-.01	.10**
pe	.07**	.07**	.05*	.02	.02
pp	.00	.06**	.04	.08**	-.04
pr	.07**	.12**	.04*	.02	-.06**
qm	-.06**	-.05*	-.04	-.06**	.08**
sr	.07**	-.14**	-.06**	-.04	-.14**
sl	-.06**	.06**	-.05*	.02	.10**
sw	-.01	.00	-.14**	-.11**	.08**
tt	-.05**	.10**	-.04*	-.05*	.09**
wc	-.01	.02	.02	-.02	.06**
we	.06**	.07**	.04*	.01	.04
yu	-.01	.03	-.06**	-.04*	.11**
mf	.05*	-.06**	.03	.06**	-.07**

Table I
FEATURES USED IN THE SYSTEM AND THEIR PEARSON'S CORRELATION COEFFICIENTS WITH PERSONALITY TRAITS AS REPORTED IN MAIRESSE ET AL. 2007. * = p SMALLER THAN .05 (WEAK CORRELATION), ** = p SMALLER THAN .01 (STRONG CORRELATION)

feature	mean	sd	min	max
ap	1	2	0	28
cm	0	1	0	19
du	0	0	0	3
el	0	0	0	3
em	0	0	0	7
im	0	0	0	3
np	0	0	0	4
ne	0	0	0	1
nb	1	4	0	64
pa	0	0	0	3
pe	0	0	0	2
pp	1	2	0	32
pr	0	0	0	8
qm	0	0	0	3
sr	0	0	0	4
sl	6	6	0	71
sw	0	0	0	1
tt	0.971	0.048	0.706	1
wc	7	7	1	79
we	0	0	0	2
yu	0	0	0	2
mf	101264	87192	68	567704

Table II
SUMMARY OF THE BEHAVIOR OF FEATURES ASSOCIATED TO PERSONALITY TRAITS IN THE DATASET.

then the system increases the score of that personality trait. If a sentence shows a feature whose frequency is higher than mean plus standard deviation and it correlates negatively with one personality trait, the system decrease the score of that personality trait. Then numerical values are turned into nominal ones ("y","n" and "o") simply checking if a value is positive, negative or it is zero. In the end the majority class of each personality trait is calculated for each user and the resulting string is taken as the user's personality model.

D. Evaluation of Personality Models

The evaluation method is based on the assumption that one user has one and only one personality and that this personality emerges at different degrees from user's posts. Hence the system evaluates the personality model comparing many posts of the same user. The drawback of this method is that we can only evaluate models for users that have more than one post in the dataset, and we have to discard all the other users.

The unsupervised system takes all the models built from the posts of a user and compares each value of the string. This evaluation method provides two measures, accuracy (a) and validity (v), defined in the formulas below:

$$a = \frac{tp + tn}{tp + tn + fp + fn} \qquad v = 1 - \frac{a}{P}$$

where P is the number of posts of one user; tp is the sum of each personality attribute matching within the same user (for example "y" and"y", "n" and "n", "o" and "o"); tn is the sum of opposite attributes within the same user ("y" and "n", "n" and "y"); fp is the sum of possible attributes turned to the balance value within the same user ("y" to "o" and "n" and "o") and fn is the sum of the zero attributes turned to positive or negative ("o" to "y" and "o" and "n"). Accuracy gives a measure of the reliability of the personality model and validity gives information about how much the model is valid for all the user's posts, in other words how much the user writes expressing the same personality traits. A low validity score means that the user shows variability in his/her writing style.

III. ANALYSIS AND DISCUSSION

We filtered out group posts (because many users with different personalities can post in a group) and kept only single users from the dataset. Most users (592) have just one post and the models obtained from those users were not considered reliable (accuracy is set to 0). Excluding the users with only one post the average accuracy is 0.631 and the average validity is 0.729. Accuracy is in line with the classification accuracies reported by Mairesse et al. 2007 for observer ratings evaluation. This fact is very encouraging because it is a clue that we developed a system that implements Mairesse's model completely automatically. The results of the frequency of personality models in the sample is reported in Table III. Below rank 7 models become more and more sparse, with a long tail of models appearing only once. Thay do not appear in Table III.

The most frequent personality type in the Italian subset of FriendFeed is represented by the model of an extravert, insecure, agreeable, organized and unimaginative person. It is interesting to note that the features "insecure" and

 ISBN: 978-1-61208-176-2

Rank	Model	Rel.Freq.
1	ynyyn	16.6%
2	ynyon	12.1%
3	onoyn	7.6 %
4	ooooo	7.6%
5	ynoyn	4.5%
6	yoooo	4.5%
7	ynooo	3.8%
8	ynoyo	3.8%
9	ynoon	3.2%
10	onyoo	3.2%
11+	*others*	33.1%

Table III
FREQUENCY OF PERSONALITY MODELS.

"unimaginative" is present in the first four positions of the ranking and that no shy people is found in the first six positions. Pearson's correlation test reveal that there is a strong (+0.79) and highly significant correlation (p-value = .0003) between the accuracy and personality model types, meaning that there are certain personality types that express strongly and reliably their personality in written language, and others that do not. Althought there is no correlation (p-value = .413) between personality and posting activity, once filtered out the long tail of users with sparse personality models, emerges that there is one personality type that produces more posts than others, that is the extravert, insecure, friendly, not particularly precise and unimaginative person (`ynyon`).

A manual look to the data reveals that there are some users (the ones with higher validity) that are focused on a topic, and sometimes this topic is clear from their username: for example "styleandthecity", or such users as "ultimora" or "cronaca24", which appear to be journalists and have a very recognizable and normalized style, but not the same personality model.

IV. CONCLUSIONS AND FUTURE WORK

In this work, we described and developed an unsupervised system for personality recognition that does not require subjects for evaluation. It exploits existing correlations between language cues and personality traits providing accuracy and validity as evaluation measures. We showed that it is possible to extract personality information from SNSs in an unsupervised way with acceptable accuracy with a process that is completely automatic. The results reported here show that the distribution of personality models in SNSs has a high peak of people sharing the same personality traits and a long tail of people with a unique personality model. Results also show that validity is a good measure of the recognizability of the style of a user.

In the future, we would like to improve the system exploiting different correlation sets. We would also like to sample and automatically annotate large corpora of Social Network data in order to facilitate the research in this field.

ACKNOWLEDGEMENT

This work has been realized also thanks to the support from the Provincia autonoma di Trento and the Fondazione Cassa di Risparmio di Trento e Rovereto.

REFERENCES

[1] Amichai-Hamburger, Y. and Vinitzky, G. Social network use and personality. In *Computers in Human Behavior*. 26(6). pp. 1289–1295. 2010.

[2] Boyd, D. and Ellison, N. Social Network Sites: Definition, history, and scholarship. In *Journal of Computer-Mediated Communication* 13(1). pp. 210–230. 2007.

[3] Celli, F. and Di Lascio and F.M.L. and Magnani, M. and Pacelli, and B., Rossi, L. *Social Network Data and Practices: the case of Friendfeed*. Advances in Social Computing, pp. 346–353. Series: Lecture Notes in Computer Science, Springer, Berlin. 2010.

[4] Golbeck, J. and Robles, C., and Turner, K. Predicting Personality with Social Media. In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*, pp. 253–262. 2011.

[5] Goldberg, L., R. The Development of Markers for the Big Five factor Structure. In *Psychological Assessment*, 4(1). pp. 26–42. 1992.

[6] Mairesse, F., and Walker, M. Words mark the nerds: computational models of personality recognition through language. In: *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. pp. 543-548. 2006.

[7] Mairesse, F. and Walker, M. A. and Mehl, M. R., and Moore, R, K. Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. In *Journal of Artificial intelligence Research*, 30. pp. 457–500. 2007.

[8] Norman, W., T. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality rating. In *Journal of Abnormal and Social Psychology*, 66. pp. 574–583. 1963.

[9] Oberlander, J., and Nowson, S. Whose thumb is it anyway? classifying author personality from weblog text. In *Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics ACL*. pp. 627–634. 2006.

[10] Quercia, D. and Kosinski, M. and Stillwell, D., and Crowcroft, J. Our Twitter Profiles, Our Selves: Predicting Personality with Twitter. In *Proceedings of SocialCom2011*. pp. 180–185. 2011

[11] Rossini Favretti R. and Tamburini F., and De Santis C. CORIS/CODIS: A corpus of written Italian based on a defined and a dynamic model. In A Rainbow of Corpora: Corpus Linguistics and the Languages of the World, Wilson, A., Rayson, P. and McEnery, T. (eds.), Lincom-Europa, Munich. pp: 27–38. 2002.

 ISBN: 978-1-61208-176-2

Cellular Automata: Simulations Using Matlab

Stavros Athanassopoulos[1,2], Christos Kaklamanis[1,2], Gerasimos Kalfoutzos[1], Evi Papaioannou[1,2]
[1]Dept. of Computer Engineering and Informatics, University of Patras
[2]Computer Technology Institute and Press "Diophantus"
Patras University Campus, Building B, GR26504, Rion, Greece
e-mail: {athanaso, kakl, kalfount, papaioan}@ceid.upatras.gr

***Abstract*—This paper presents a series of implementations of cellular automata rules using the Matlab programming environment. A cellular automaton is a decentralized computing model providing an excellent platform for performing complex computations with the help of only local information. Matlab is a numerical interactive computing environment and a high-level language with users coming from various backgrounds of engineering, science, and economics that enables performing computationally intensive tasks faster than with traditional programming languages (such as C, C++, and Fortran). Our objective has been to investigate and exploit the potential of Matlab, which is simple mathematical programming environment that does not require specific programming skills, regarding the understanding and the efficient simulation of complex patterns, arising in nature and across several scientific fields, captured by simple cellular automata structures. We have implemented several cellular automata rules from the recent literature; herein we present indicative cases of practical interest: the forest fire probabilistic rule, the sand pile rule, the ant rule, the traffic jam rule as well as the well-known "Game of Life". Our work indicates that Matlab is indeed an appropriate environment for developing simulations for cellular automata models.**

Keywords-cellular automata; simulation; Matlab.

I. CELLULAR AUTOMATA

A cellular automaton (CA) is an idealization of a physical system in which space and time are discrete and the physical quantities take only a finite set of values. Informally, a cellular automaton is a lattice of cells, each of which may be in a predetermined number of discrete states (a formal definition can be found in [7]). A neighborhood relation is defined over this lattice, indicating for each cell which cells are considered to be its neighbors during state updates. Time is also discrete; in each time step, every cell updates its state using a transition rule that takes as input the states of all cells in its neighborhood (which usually includes the cell itself). All cells in the cellular automaton are synchronously updated. At time t = 0 the initial state of the cellular automaton must be defined; then repeated synchronous application of the transition function to all cells in the lattice will lead to the deterministic evolution of the cellular automaton over time. Many variations of this basic model exist: CA can be of arbitrary dimension, although one-dimensional and two-dimensional CA have received special attention in the literature. CA can be infinite or finite. Finite CA can have periodic boundaries (e.g., the opposite ends of a one-dimensional finite CA are joined together so the whole forms a ring). Updates can be synchronous or asynchronous. Transition rules can be deterministic or stochastic. Many other variations exist; those mentioned above are some of the most typical ones.

The concept of cellular automata was initiated in the early 1950's by John Von Neumann and Stan Ulam [18]. Von Neumann was interested in their use for modelling self-reproduction and showed that a CA can be universal. He devised a CA, each cell of which has a state space of 29 states, and showed that it can execute any computable operation. However, Von Neumann rules, due to their complexity, were never implemented on a computer. Von Neumann's research raised a dichotomy in CA research. On one hand, it was proven that a decentralized machine can be designed to simulate any arbitrary function. On the other hand, this machine (CA) can become as complex as the function it is intended to simulate.

Cellular automata have received extensive academic study into their fundamental characteristics and capabilities and have been applied successfully to the modelling of natural phenomena and complex systems [1], [3], [4], [13], [17], [24], [23]. Based on the theoretical concept of universality, researchers have tried to develop simpler and more practical architectures of CA that can be used to model widely divergent application areas. In the 1970, the mathematician John Conway proposed the (now famous) Game of Life [10], which received widespread interest among researchers. Since the beginning of the 80's, Stephen Wolfram has studied in much detail a family of simple one-dimensional cellular automata rules (known as Wolfram rules [24]) and has showed that even these simplest rules are capable of emulating complex behavior. Other applications include, but are not limited to, theoretical biology [2], game theory [19], and non-equilibrium thermodynamics [15].

The rest of the paper is structured as follows: Section II includes a brief description of Matlab as well as main reasons that motivated us for using it in our simulations. Simulations are presented in Section III. Section IV includes conclusion and plans for future work on cellular automata simulations using Matlab.

 ISBN: 978-1-61208-176-2

II. MATLAB

MATLAB is a numerical computing environment and fourth-generation programming language which allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, Java, and Fortran. Although it was intended primarily for numerical computing, it also allows symbolic computing, graphical multi-domain simulation and model-based design for dynamic and embedded systems. It has been widely used in academia and industry by users coming from various backgrounds of engineering, science and economics. MATLAB was first adopted by researchers and practitioners in control engineering, and quickly spread to many other domains. It is now also used in education, in particular the teaching of linear algebra and numerical analysis, and is very popular amongst scientists involved in image processing [16].

Why we used Matlab for our simulations? Existing implementations of cellular automata have been developed using Java and C/C++. This selection has been supported by the graphical interface these programming languages offer as well as by their strict object-oriented programming nature. In this way, implementation of cellular automata can be a very efficient and effective development task. For our study, Matlab offers simplicity coupled with power; this mainly motivated us to use it for the implementation/simulation of cellular automata, i.e., of simple structures that can, however, model complex behavior and real-world patterns. Matlab neither requires nor focuses on particular programming skills; on the contrary, it provides an efficient tool for the researcher/user to simulate simple models without focusing on programming and easily conceive such complex patterns in practice – not only through some mathematically defined function (however, using appropriate toolboxes, Matlab code can be converted – if needed – to C/C++ code).

More specifically, cellular automata can be implemented using matrices of one or several dimensions. Matlab makes a quite appropriate environment since it offers a wide range of operations and functions particularly working on matrices. Moreover, the status of network cells can be easily represented using function surf(), while necessary diagrams and graphical representations can be produced - almost directly - using function plot(). Using Matlab only a single file per cellular automaton (i.e., per algorithm) is needed; this provides high flexibility in the experimentation and simplicity in the code execution process. Furthermore, syntax is simpler (than in involved programming languages) thus directly reflecting the simplicity of the rules according to which automaton cell status is altered. Such technicalities could be of high importance when it comes to communities of researchers not familiar with programming languages: they could easily deploy their model and see its behavior without having to spend extra resources for becoming programming experts. Of course, Matlab is a rather slow environment and Matlab programs require more computational power compared to Java or C++; this could be a drawback if our algorithms were to be used as parts of intense resource-requiring applications.

III. OUR SIMULATIONS

As already stated, the question that motivated our work is the following: Matlab is a "simple" programming environment that does not require a researcher/student to be a programming-expert to use it. Cellular automata can capture, via a small set of simple rules, very complex phenomena from the real world. Is Matlab efficient for simulations involving cellular automata?

We have implemented several CA rules from the recent literature: the Wolfram's 184 rule, rules for probabilistic cellular automata, the Q2R rule, the annealing rule, the HPP rule, the sand pile rule, the ant rule, the traffic jam rule, the solid body motion rule, the "*Game of Life*". Detailed description of these rules can be found in [7].

Herein, we present in detail five indicative cases of practical interest we simulated (and used for teaching purposes in the Theory of Computation lab of our department): Probabilistic Cellular Automata rules for forest fire models, the Sand Pile rule, the Ant rule, the Traffic Jam rule and the John Conway's Game of Life.

Matlab Version 7.0.0.19920 (R14) has been used for implementation. Simulations have been executed on a system using an Intel Core i3 530 processor (2.93GHz, 6144MB DDR3 RAM), running Windows 7 Premium 32-bit operating system. For the graphical representation of the behavior of simulated models, function surf() has been used (full size figures can be found at [25]).

A. Implementation of a probabilistic rule for Burning Forest

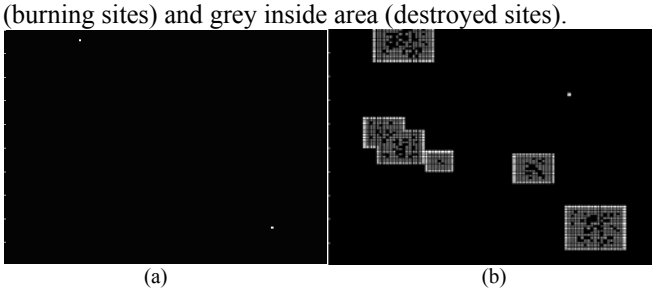
Probabilistic Cellular Automata (PCA) are ordinary cellular automata where different rules can be applied at each cell according to some probability [24]. An interesting and simple example of a PCA model is a probabilistic rule for Burning Forest. The cellular automaton used for simulation uses a (nxn) grid, representing the forest, and a Moore neighborhood. Cells correspond to trees and can be in one of the following three states: green tree (1), empty site (2), burning tree (3). Initially, all cells are in state (1) (i.e., contain a green tree). Cell states are updated according to the following rules presented in detail in [5], [8]:

- A burning tree becomes an empty site.
- A green tree becomes a burning tree if at least one of its nearest neighbors is burning.
- At an empty site, a tree grows with probability p.
- A tree without a burning nearest neighbor becomes a burning tree in one time step with probability f.

 ISBN: 978-1-61208-176-2

At each time step, every cell is assigned a new random value (in [0,1]) for fire (f) and birth (p) probability. A green tree becomes a burning tree when f is greater than a threshold value set to 0.001. A new tree grows in an empty site when p is greater than a threshold value set to 0.1. These threshold values for f and p, once set remain the same throughout a single execution. Threshold value for f has been chosen to be sufficiently small so that in a large grid only few fires can start. Threshold value for p has been chosen to be greater than this for f so that new trees can grow and simulation can continue.

The following figures show instances of the simulation using a grid of size 200x200. In the beginning (Fig. 1a) two fires (white areas) have started in the forest (black area). Fire starts spreading among green trees, leaving empty sites behind (grey areas) (Fig. 1b). The fire spreading pattern looks like growing circular discs with a white outline (burning sites) and grey inside area (destroyed sites).



(a) (b)

Fig. 1: Two fires have started in the forest (white sites) (a). The fire is spreading among green trees, turning them to empty sites (b).

B. Implementation of the Sand Pile rule

The physics of granular materials has recently attracted CA-related research interest. It is possible to devise a simple cellular automaton rule to model basic piling and toppling of particles like sand grains [7]. The idea is that grains can stack on top of each other if this arrangement is stable. Of course, real grains do not stand on a regular lattice and the stability criteria are expected to depend on the shape of each grain. Despite the microscopic complexity, the result is sand piles that are too high to topple.

Toppling mechanisms can be captured by the following cellular automaton rule: a grain is stable if there is a grain underneath and two other grains preventing it falling to the left or right (Fig. 2). Assuming a Moore neighborhood, the rule implies that a central grain will be at rest if the south-west, south and south-east neighbors are occupied. Otherwise, the grain topples downwards to the nearest empty cell.

Figure 2: The top grain will not move.

The cellular automaton used for simulation uses a (nxn) grid and a Margolus neighborhood which gives a simple way to deal with the synchronous motion of all particles [20]. Informally, when Margolus neighborhood is used, the lattice is divided in disjoint blocks of size 2x2; each block moves down and to the right with the next generation, and then moves back [21]. Cells can be in one of the following two states: grain of sand (1), empty cell (0). Initially, sand grains are placed randomly on the grid (no additional grains appear during the evolution of the cellular automaton). Cell states are updated according to the following rule [7], which is also presented graphically in Fig. 3:

Current state	1000	0100	1010	1001	0110	0101	1110	1101	1100 (p)	1100 (1-p)
Next state	0010	0001	0011	0011	0011	0011	1011	0111	0011	11100

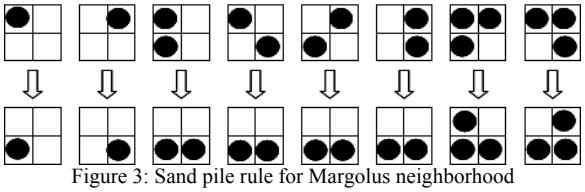


Figure 3: Sand pile rule for Margolus neighborhood

The configuration in which the upper part of a block is occupied by two particles while the lower part is empty, is not listed in the above image, although it certainly yields some toppling. When this configuration occurs, we adopted the probabilistic evolution rule shown in Fig. 4 in order to produce a more realistic behavior: some friction may be present between grains and some arches may appear to delay collapse. Of course, the toppling of other configurations could also be controlled by a random choice.

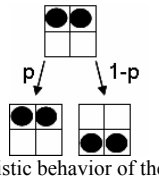
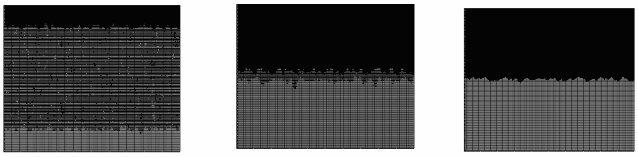


Figure 4: Probabilistic behavior of the sand pile rule [7]

In this simulation, p has been set to 0.5, i.e., two neighboring grains can equiprobably either fall (filling the cells bellow them) or remain at rest.

Fig. 5a, 5b and 5c show simulation instances. Initially, all grains are falling, except those at the bottom which remain at rest. The Margolus neighborhood does not affect grains at the grid boundaries, so they also remain at rest. The sand pile is growing and the number of falling grains decreases (Fig. 5b). The simulation terminates when there are no more grains to fall (Fig. 5c).



(a) (b) (c)

Figure 5: The initial state of the lattice (a). The growing sand pile due to falling grains (b). Finally, a sand pile is created (c)

C. Implementation of the Ant rule

Langton's Ant [13, 14] follows extremely simple rules and initially appears to behave chaotically, however after a

certain number of steps a recurring pattern emerges. Langton's Ant models the true behavior of ants in nature: a moving ant tends to leave pheromones behind it. All other ants moving in the same area can sense that substance and follow the motion of the first ant.

The rule simulates the following idea: an ant sits on a cell of a grid where all other cells are initially empty. It moves into a neighboring cell and does one of two things, based on the color of the cell:

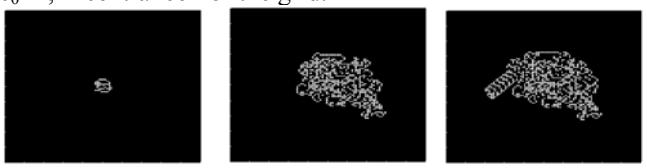
– If the square is white, it turns 90 degrees to the left and colors the square grey
– If the square is grey, it turns 90 degrees to the right and colors the square white

The movement is continued in the same fashion, ad infinitum. The interesting thing about this is that after a fixed number of steps, the ant builds a highway and hotfoots it into infinity. The motion of the ant in this highway is not linear; it rather looks like the pattern of operation of a sewing machine. Although the ant rule seems to be very simple, it drives the ant to a chaotic state. This feature also shows the power of modeling systems with cellular automata: even though the cellular automata rules are very simple, they can implement very complex behaviors.

The cellular automaton used for simulation uses a (nxn) grid and a Von Neumann neighborhood; a von Neumann neighborhood is composed of the four cells orthogonally surrounding a central cell on a two-dimensional square lattice [12]. A cyclic neighbourhood has been used for cells at the lattice boundaries: when an ant reaches the lattice boundaries, it returns to the lattice simulating the existence of a second ant. Cells can be in one of two states: ant (1), empty cell (0). Initially, all cells are empty (state 0) apart from one cell (state 1) which contains the ant. Cell states are updated according to the following rules:

– $n_i (r + c_i, t + 1) = \mu\, n_{i-1} (r, t) + (1 - \mu)\, n_{i+1} (r, t)$
– $\mu (r, t + 1) = \mu (r, t) \oplus n_1(r, t) \oplus n_2(r, t) \oplus n_3(r, t) \oplus n_4(r, t)$

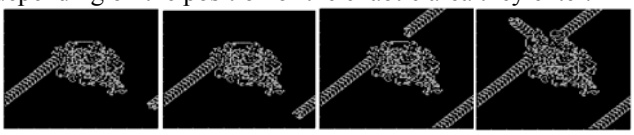
where n_i: new state, r: current cell, c_i: current direction, t: current time step, μ: cell color (1=white, 0=black). Initially, c_0=4, r=central cell of the grid.



(a) (b) (c)

Figure 6: (a) The ant starts its journey from the centre of the lattice. (b) Chaotic situation due to the ant movement. (c) Ant's highway.

In our simulation, an ant starts its journey from the central cell of a 100x100 grid (Fig. 6a). All cells are initially black; the ant turns them white as it moves over them. After approximately 7000 steps, the ant is trapped in a chaotic situation (Fig. 6b). After approximately 10000 steps, the ant creates its way out of the chaotic situation, building its highway and moving away from its initial position (Fig. 6c). As soon as the ant reaches the lattice boundary it returns to the lattice from a different position as a "second" ant, which has just entered the area. This second ant continues moving on the highway, just like the first one (Fig. 7a), moves towards the chaotic area (created by the first ant) (Fig. 7b) and starts moving irregularly (Fig. 7c). The "second" ant "senses the pheromones" of the first ant and escapes the chaotic situation faster than the previous ant. Ants can either create their own highways (Fig. 7d) or follow existing ones depending on the position of the chaotic area they enter.



(a) (b) (c) (d)

Figure 7: the movement of a second ant

D. Implementation of the Traffic Jam rule

Cellular automata models for road traffic have received a great deal of interest. One-dimensional models for single-lane car motions are quite simple and elegant [6]. The road is represented as a line of cells: each cell is either occupied by a vehicle or not. All cars travel in the same direction. Their positions are updated synchronously, in successive iterations (discrete time steps). During the motion, each car can be at rest or jump to the nearest-neighbor site, along the direction of motion. The rule is that a car moves only if its destination cell is empty. This means that the drivers are short-sighted and do not know whether the car in front will move or whether it is also blocked by another car. Therefore the state of each cell s_i is entirely determined by the occupancy of the cell itself and its two nearest neighbors s_{i-1} and s_{i+1}. The motion rule can be summarized in the following table, where all eight possible configurations $(s_{i-1}\ s_i\ s_{i+1})_t \rightarrow (s_i)_{t+1}$ are given [6]:

111	110	101	100	011	010	001	000
1	0	1	1	1	0	0	0

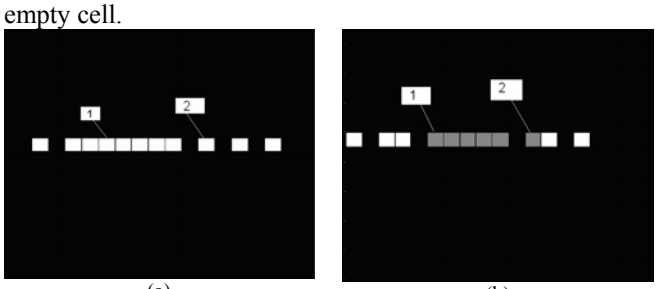
This simple dynamics captures an interesting feature of real car motion: traffic congestion. This cellular automaton rule turns out to be Wolfram's rule 184 [6].

The cellular automaton used for simulation uses a line and a one-dimensional neighborhood. Cells can be in one of three states: empty cell (0), stopped car (1), moving car (2). Initially, cars are placed randomly in line cells. Cell states are updated according to the following rule:

– $n_i(t+1)=n_i^{in}(t)(1-n_i(t))+ n_i(t)\, n_i^{out}(t)$,

where $n_i(t)$ denotes the car occupation number (n_i=0: free site, n_i=1: a car is present at site i). $n_i^{in}(t)$ denotes the state of the source cell, i.e., that from which a car may move to cell i. Similarly, $n_i^{out}(t)$ indicates the state of the destination cell, i.e., that the car at site i would like to move to. The rule implies that the next state of cell i is 1 if a car is currently present and the next cell is occupied, or if no car is currently present and a car is arriving.

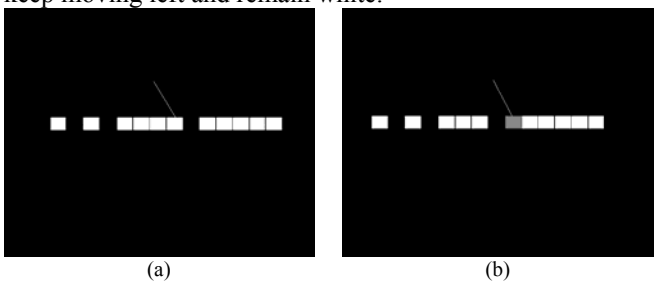
A car is moving or not according to its "speed", a variable taking random values in [0,1] that change in each time step. If a car has a "speed" lower than a threshold value set to 0.05, then it stops for one time step. When a car reaches the leftmost cell of its row, it is injected in the rightmost cell of the lattice in the same row and keeps moving in loops. Fig. 8a shows a normal traffic instance where all the cars are moving from right to left by one cell per step. White cars are moving; grey cars have stopped. In Fig. 8b, cars 1 and 2 stop. When car 1 stops, all following cars also stop (since there are no empty cells between them) and turn grey. Cars in front of car 1 keep moving left because no preceding car has stopped. When car 2 stops, there is an empty cell behind it. This is why the following cars remain white and keep moving left, covering every empty cell.



(a) (b)

Figure 8: Normal traffic (a) and traffic with cars that are not moving (b).

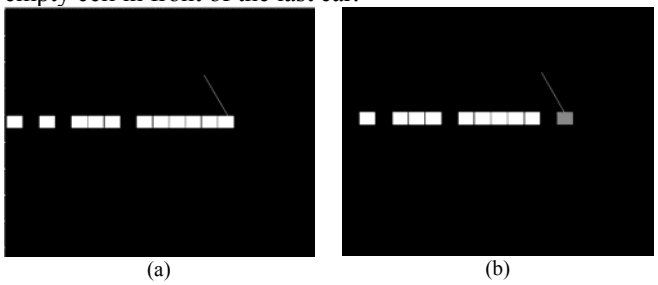
In Fig. 9a, another instance of normal traffic is shown. The car pointed by the arrow stops and becomes grey (Fig. 9b). There is an empty cell behind it, so all cars that follow keep moving left and remain white.



(a) (b)

Figure 9: Normal traffic (a) and then a car stops (b)

Finally, the last car (Fig. 10a) stops (and becomes grey in Fig. 10b). All other cars keep moving left leaving an empty cell in front of the last car.



(a) (b)

Figures 10: Normal traffic (a) and then a car stops (b)

E. John Conway's Game of Life

The Game of Life is a cellular automaton devised by the British mathematician John Horton Conway in 1970 [10]. The game is a zero-player game, meaning that its evolution is determined by its initial state (called pattern), requiring no further input. One interacts with the Game of Life by creating an initial configuration and observing how it evolves. The universe of the Game of Life is an infinite two-dimensional orthogonal grid of square cells, each of which is in one of two possible states, alive (white-1) or dead (black-0). Every cell interacts with its eight neighbours (Moore neighbourhood), which are the cells that are horizontally, vertically, or diagonally adjacent. Cell states are updated according to the following rule:

- Any live cell with fewer than two live neighbours dies, as if caused by under population.
- Any live cell with two or three live neighbours lives on to the next generation.
- Any live cell with more than three live neighbours dies, as if by overcrowding.
- Any cell with exactly three live neighbours becomes a live cell, as if by reproduction.

The initial pattern placed in the middle cells of the grid constitutes the seed of the system. The first generation is created by applying the above rules simultaneously to every cell in the seed-births and by deaths occurring simultaneously, and the discrete moment at which this happens is sometimes called a tick; in other words, each generation is a pure function of the preceding one. The rules continue to be applied repeatedly to create further generations.

Fig. 11 shows simulation snapshots for 6 different initial patterns: cell row (Fig. 11a), glinder (Fig. 11b), small explorer (Fig. 11c), explorer (Fig. 11d), lightweight spaceship (Fig. 11e), tumbler (Fig. 11f).

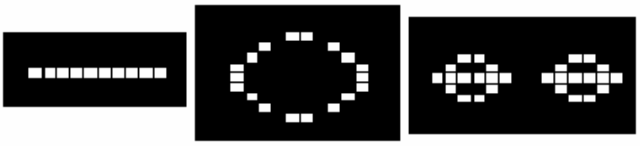


Figure 11a: Cell Row

Figure 11b: Glinder

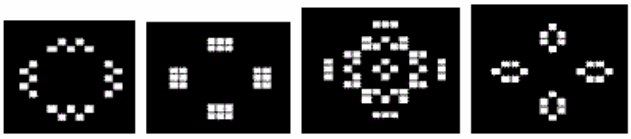


Figure 11c: Small explorer

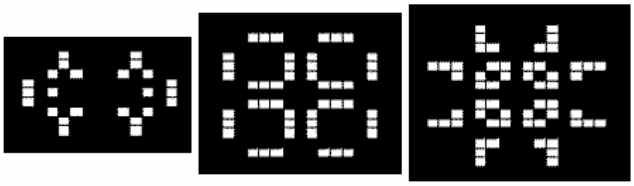


Figure 11d: Explorer

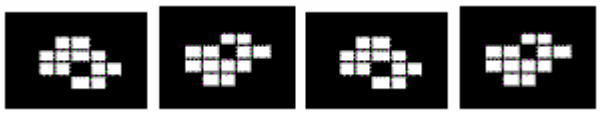


Figure 11e: Lightweight spaceship

Figure 11f: Tumbler

IV. CONCLUSION AND FUTURE WORK

We have simulated several popular cellular automata rules of practical interest using Matlab. Our simulations yield evolution patterns in accordance with those expected from corresponding rules and similar to those obtained so far using Java or C/C++.

Our work indicates that Matlab is indeed an appropriate environment for developing compact code for simulations involving cellular automata, even though it does not always guarantee high simulation speeds. It does not require specific programming skills and therefore it offers the flexibility to non-programming-expert researchers and/or students to experiment and understand in practice complex patterns captured by simple cellular automata structures.

Our current ongoing work investigates the potential of Matlab for simulations involving cellular automata for problems related to energy-efficient communication in Wireless Sensor Networks.

ACKNOWLEDGMENT

This work has been partially supported by EU under the ICT-2010-258307 project EULER and by EU and the Greek Ministry of Education, Lifelong Learning and Religious Affairs under the project “Digital School” (296441).

REFERENCES

[1] S. Bandini. Guest Editorial - Cellular Automata. Future Generation Computer Systems, 18:v-vi, August 2002.

[2] M. Boerlijst and P. Hogeweg. Spiral wave structure in pre-biotic evolution: hypercycles stable against parasites. Physica D, 48(1):17–28, 1991.

[3] A. W. Burks. Essays on Cellular Automata. Technical Report, Univ. of Illinois, Urbana, 1970.

[4] P. Pal Chaudhuri, D. R. Chowdhury, S. Nandi, and S. Chatterjee. Additive Cellular Automata - Theory and Applications, volume 1. IEEE Computer Society Press, CA, USA, ISBN 0-8186-7717-1, 1997.

[5] P. Bak, K. Chen, C. Tang. A forest-fire model and some thoughts on turbulence. Physics Letters A, Vol. 147, Issues 5-6, pp. 297-300, 1990.

[6] B. Chopard, P. O. Luthi, and P-A. Queloz. Cellular Automata Model of Car Traffic in a Two-Dimensional Street Network. Journal of Physics A: Mathematical and General, 29, pp. 2325–2336, 1996.

[7] B. Chopard and M. Droz. Cellular Automata Modeling of Physical Systems, Cambridge University Press, 1998. ISBN 0-521-46168-5.

[8] B. Drossel, F. Schwabl. Self-organized critical forest-fire model. Physical Review Letters, Vol. 69, Issue 11, pp. 1629–1632. 1992.

[9] N. Ganguly, B. K. Sikdar, A. Deutsch, G. Canright, and P. Pal Chaudhuri. A survey on cellular automata. Technical report, Centre for High Performance Computing, Dresden University of Technology, December 2003.

[10] M. Gardner Mathematical Games - The fantastic combinations of John Conway's new solitaire game "life". Scientific American, 223. pp. 120-123, 1970. ISBN 0894540017.

[11] R. Goering. Matlab edges closer to electronic design automation world. EE Times, 10/04/2004 (http://www.eetimes.com/electronics-news/4050334/Matlab-edges-closer-to-electronic-design-automation-world).

[12] L. Gray. A Mathematician Looks at Wolfram's New Kind of Science. Not. Amer. Math. Soc. 50, 200-211, 2003.

[13] C. G. Langton. Self-reproduction in cellular automata. Physica D: Nonlinear Phenomena, Volume 10, Issues 1-2, pp. 135-144, 1984. ISSN 0167-2789, 10.1016/0167-2789(84)90256-2.

[14] C. G. Langton. Studying artificial life with cellular automata. Physica D: Nonlinear Phenomena 22 (1-3): 120–149, 1986.

[15] M. Markus B. Hess. Isotropic cellular automaton for modelling excitable media. Nature, 347(6288):56–58, 1990.

[16] C. Moler. The Origins of MATLAB. December 2004. Retrieved April 15, 2007.

[17] M. Mitchell, P. T. Hraber, and J. P. Crutcheld. Revisiting the Egde of Chaos: Evolving Cellular Automata to Perform Computations. Complex Systems, 7, pp. 89-130, 1993.

[18] J. V. Neumann. The Theory of Self-Reproducing Automata. A. W. Burks (ed), Univ. of Illinois Press, Urbana and London, 1966.

[19] M. Nowak and R. May. Evolutionary games and spatial chaos. Nature, 359(6398):826–829, 1992.

[20] J. Schiff. 4.2.1 Partitioning Cellular Automata. Cellular Automata: A Discrete View of the World, Wiley, pp. 115–116, 2008.

[21] T. Toffoli, N. Margolus. II.12 The Margolus neighborhood. Cellular Automata Machines: A New Environment for Modeling, MIT Press, pp. 119–138, 1987.

[22] S. Wolfram. A New Kind of Science. Champaign, IL: Wolfram Media, pp. 29-30, 52, 59, 317, and p. 871, 2002.

[23] S. Wolfram. Cellular Automata and Complexity. World Scientific, Singapore, 1994. ISBN 9971-50-124-4 pbk.

[24] S. Wolfram. Theory and Applications of Cellular Automata. World Scientific, Singapore, 1986. ISBN 9971-50-124-4 pbk.

[25] http://www.ceid.upatras.gr/papaioan/CA/figs/index.html, November 15, 2011.

 ISBN: 978-1-61208-176-2

Fast Polynomial Approximation Acceleration on the GPU

Lumír Janošek
Department of Computer Science
VSB-Technical University of Ostrava
Ostrava, Czech Republic
Email: lumir.janosek.st@vsb.cz

Martin Němec
Department of Computer Science
VSB-Technical University of Ostrava
Ostrava, Czech Republic
Email: martin.nemec@vsb.cz

***Abstract*—This article presents the possibility of parallelization of calculating polynomial approximations with large data inputs on GPU using NVIDIA CUDA architecture. Parallel implementation on the GPU is compared to the single thread CPU implementation. Despite the enormous computing power of today's graphics cards there is still a problem with the speed of data transfer to GPU. The article is mainly focused on the implementation of some ways of transferring data from memory into GPU memory. The aim is to show what method is suitable for a large amount of data being processed and what for the lesser amount of data. Afterwards performance characteristics of the implementation of the CPU and GPU are matched.**

Keywords-GPU; CUDA; Direct Memory Access; Parallel Reduction; Approximation.

I. INTRODUCTION

This article is focused on the application of a parallel approach to the implementation of the polynomial approximation of the k-th degree and its comparison with conventional single thread approach. Polynomial approximation model is widely used in practice. The statistics commonly use the basic model of approximation of 1th degree - a linear approximation, in other statistics called the linear regression.

Nowadays it is possible to create a massively parallelized applications using modern GPUs (Graphics Processing Unit) that enable the distribution of calculations among tens of multiprocessors of graphic cards. The problem still remains the need to transfer data between the CPU (Central Processing Unit) and GPU. This can become a limiting factor in performance when the time needed to transfer data between memory and GPU memory, the host system plus the time the GPU processes data exceeds the time after, which the same data can be handled by the CPU. But there are ways to at least partially eliminate this lack of trying.

In this article will be shown how to implement polynomial approximation using the GPU parallel computing architecture of NVIDIA CUDA (Compute Unified Device Architecture), which provides a significant increase of computing power [1]. The parallel implementation is compared with single threaded CPU implementation. Performance results of both implementations are compared with each other and show the differences between the parallel implementation approach and common single threaded approach for certain volume of data. By comparison of these two approaches it can be seen for how much data is suitable for the parallel approach and for how much it is already inappropriate. A substantial part of the implementation is a comparison of the chosen methods of copying data from RAM (Random-Access Memory) to graphics card memory, and especially the methods of allocating this memory. Three methods are compared: the allocation of pageable memory, the allocation of page-locked memory (also known as Pinned memory), and the allocation of memory mapped into the address space of the CUDA [2].

A common approach is the method of allocation and data transfer, when the input data are placed in pageable memory and from this memory are then transmitted by conventional copying approach into graphics card memory. The allocation of page-locked memory when copying data allows the GPU to use DMA (Direct Memory Access). Mapping memory allocation into the address space of the CUDA is a special case that allows to read data stored in RAM directly from the GPU.

This paper is structured as follows. First, some mathematical background related to polynomial approximations is presented. Next, a description of the implemented memory approaches and a description of the implementation of a parallel reduction are presented. Lastly, results and conclusion are presented.

II. MATHEMATICAL MODEL OF APPROXIMATION

Consider a set of points with coordinates x_i z $\mathbb{R}^d$, where $i \in \{1, \cdots, n\}$. The aim of the approximation data problem is to find the function $f(x)$ in the general case, which best approximates the scalar value f_i at point x_i. The result, using the least squares method, is a function $f(x)$ such that the distance between scalar data values f_i and functional values $f(x_i)$ is as small as possible [3]. Least squares method based linear approximation in its simplest application, which approximates an input data by linear function in the form of:

$$f : b_0 + x \cdot b_1, \tag{1}$$

 ISBN: 978-1-61208-176-2

where the sum of squares has the form:

$$\psi(b_0, b_1) := \sum_{i=1}^{n} [f(x_i) - f_i]^2 \quad (2)$$

Minimum of sum of squares then we found as:

$$\frac{\partial \psi}{\partial b_0} = 0 \qquad \frac{\partial \psi}{\partial b_1} = 0$$

By adjusting the obtained:

$$\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix} \quad (3)$$

Members of the vector of the right side b_0 and b_1 are coefficients of the polynomial approximation (1). Input data of the algorithm are represented by a set of vectors (pairs) of $\mathbb{R}^2$. For input data it is sufficient to calculate the four sums (vector of sums):

$$V_\Sigma = \left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} y_i, \sum_{i=1}^{n} x_i y_i, \sum_{i=1}^{n} x_i^2 \right) \quad (4)$$

The results of these sums are then just put back into the system of equations (3). By solving it, we get the sought coefficients of b_0 and b_1 approximation polynomial (1).

A. Polynomial approximation

A special case of linear model approximation is polynomial approximation. It is an approximation by polynomial of k-th degree. Using the procedure for calculating the linear approximation it is possible to express polynomial approximations formula as a set of [4]:

$$b = A^{-1} Y \quad (5)$$

where

$$A = \begin{pmatrix} n & \sum_{i=1}^{n} x_i & \cdots & \sum_{i=1}^{n} x_i^k \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \cdots & \sum_{i=1}^{n} x_i^{k+1} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^{n} x_i^k & \sum_{i=1}^{n} x_i^{k+1} & \cdots & \sum_{i=1}^{n} x_i^{2k} \end{pmatrix}$$

$$b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix}, Y = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \cdots \\ \sum_{i=1}^{n} x_i^k y_i \end{pmatrix}$$

The solution of this system of equations is a vector b, the individual members of which $b_0 \cdots b_k$ represent the coefficients of the approximation polynomial. Using the system of equations (5) it is possible to derive a vector of sum for any polynomial k-th degree as in the case of linear approximation.

III. Implementation of memory approaches

The algorithm for calculating polynomial approximation, which was described in the previous section, was implemented using the CUDA architecture of NVIDIA. The implementation was designed for processing large amounts of data represented as a set of vectors of $\mathbb{R}^2$. Parallel implementation of polynomial approximations on GPU is compared with single threaded implementation on CPU. During implementation, the goal was due to the large amount of input data, to optimize data flow between RAM and GPU memory.

Given the size of the input data set, several approaches to copy data from RAM to graphics card memory (global memory), or to access data from the GPU were compared. Three approaches were compared: normal approach with pageable memory allocation, the allocation of page-locked memory (also known as Pinned memory), and the allocation of page-lock memory mapped into address space of the CUDA.

The actual calculation of the polynomial approximation was implemented in part on the GPU and in part on the CPU. For comparison approximations of 1th degree, 2th degree and 3th degree were implemented. A parallel approach was used for calculating the vector of sums (4) by using a the parallel reduction algorithm. Calculation of the resulting coefficient b_0 and b_1 of the approximation polynomial is then completed on the CPU.

A. Pageable system buffer and page-locked memory

A common approach to transfer data from RAM to the global memory was compared with direct access of the GPU to RAM when copying data, otherwise the DMA (Direct Memory Access). The disadvantage of the common approach is double copying of transferred data. Data are transmitted in the first step from pageable memory (pageable system buffer) to the page-locked memory, and then from this page-locked memory to the GPU memory. By direct allocation of a data buffer in the page-locked memory extra data copying can be avoided. By allocation page-locked memory the operating system guarantees that this memory is not paged to disk, thus ensuring its place in physical memory [5]. By knowing the physical address of the buffer in the memory, GPU can copy data to the global memory direct memory access - DMA.

B. Memory mapping into address space of the CUDA

Another approach to access data in the RAM from the GPU is using direct mapping of page-lock memory into address space of the CUDA. The data are, as in the previous case, stored in memory allocated as page-locked memory, with the only difference being that this data can be accessed directly from the GPU. This eliminates the need for allocating memory block in global memory and the need to copy data into this block of memory.

 ISBN: 978-1-61208-176-2

IV. Parallel reduction

With access to parallel hardware the entire process of the sum calculation can be parallelized. If we have hundreds of threads, then each thread can contribute to the gradual calculation of the resulting sum. This approach is called parallel reduction [6]. The main idea of the parallel reduction is that each thread performs the sum of two values in the memory and then saved back. The algorithm therefore starts at the beginning with half the number of threads than the number of inputs. In each step, one thread adds the two values. In the next step the process is repeated, but with half the number of threads. The process continues until the final sum is achieved by gradual reduction. The parallel reduction algorithm is especially efficient for large data inputs.

Reduction of the vector (4) is divided between $C \cdot N_{threads}^{-1}$ blocks, where C is the count of input data (vectors of $\mathbb{R}^2$) and $N_{threads} = 256$ is the number of threads per block. The data are this way evenly divided between the individual blocks, when each block handles one subset of the input data.

Implementation of the parallel reduction of the vector of sums can be divided into three steps: 1) The first step is to copy data from global memory to the shared memory. In the shared memory the reduction of the vector of sums is subsequently made. Copying data from global memory to shared memory is implemented in the CUDA kernel by using all threads of the block, thus each thread copies always the four values that belong to one of the four sums of a vector (4). Simultaneously with copying the data into shared memory, is made the first reduction step - first add during load [6]. This leads to the reduction of the required number of blocks by half. The total number of blocks needed to run the CUDA kernel is

$$\frac{1}{2} \cdot \frac{C}{N_{threads}}$$

2) After copying the data into shared memory all the threads of block are synchronized, which ensures that no thread starts reading the shared memory until all threads finish copying the data. Then begins the process of reduction. In each iteration, one thread performs the sum of the vector of sums, which leads to a gradual reduction of input data. Before entering the next iteration, the number of threads is reduced by half. Reduction cycle ends when the number of threads reaches zero. The data inside the loop are processed in the shared memory (on-chip memory), accordingly there is no unnecessary transfer of data between global memory and GPU multiprocessor. 3) The result of each block is transferred back to global memory after the reduction. This copy process takes place before the end of kernel one of the threads. The results of the individual blocks are copied from global memory back to the RAM on the CPU. Completion of reduction, thus the sum of all results of individual blocks, is completed on the CPU. The result is a vector (4). It is then possible, without difficulty, to apply the described algorithm of a parallel reduction, with minor modifications, to the calculation of vectors of sums of approximations of higher degrees.

V. Results

The presented method for parallel calculation of linear approximation to the GPU has been implemented and tested on a graphics card GeForce 9600 GT, GeForce 9800 GTX and GeForce GTS 450 NVIDIA. The implementation was tested for various sizes of the input file in order to determine what amount of data is preferable to compute on the CPU and for how much data it is more efficient to use a parallel implementation on GPUs. Performance characteristics of both implementations were compared, the result is shown in the Fig. 1. From a comparison of the characteristics of computations on the CPU and GPU it is obvious that for a smaller amount of data it is preferable to keep the calculation of polynomial approximation on the CPU. GPU in this case is more appropriate for larger data amounts. The size of test data ranged from 12 KB to 50 MB (1365 - 6400000 input data).

As written in the previous section, three approaches of transferring data between RAM and the global memory were compared. A common approach to copy data between RAM and global memory of GPU is compared with the approach of direct access of the GPU to RAM (DMA) when copying data. This method of implementation has brought strong effect especially in an expanding volume of copied data, as seen from the Figure 2 because there is no need to copy data from pageable memory into the page-locked memory, before transferring data to the GPU global memory.

The last of the studied approaches of transferring data from memory into the GPU global memory was the use of mapping page-lock memory into address space of the CUDA. Mapping page-lock memory is especially suitable for integrated graphics processors that are built into the system chipset and usually share their memory with the CPU. In this case, using the mapping page-lock memory removes unnecessary data transfers. For discrete graphics processors, the mapping page-lock memory is only suitable just in some cases [7]. For this reason, this method also did not bring any optimization of implementation. On the other contrary, when using the mapping page-lock memory into address space of the CUDA, there was a significant downgrade in performance, see Figure 3. Below are listed the size of data transfers that have occurred during the calculation between the CPU and GPU memory.

A total number of N bytes of data was transmitted into the global memory from RAM. After completion of the calculation on the GPU back to RAM was transmitted a total of

$$\|V_\Sigma\|_2 \cdot \left(\frac{NumBlocks}{2} \cdot 4_{bytes} \right)$$

 ISBN: 978-1-61208-176-2

The total number of bytes transferred from global memory back into the RAM is equal to:

$$\|V_\Sigma\|_2 \cdot \left(\frac{1}{2} \cdot \frac{N}{4_{bytes}} \cdot \frac{1}{NumThreads} \cdot 4_{bytes}\right)$$

$$\frac{1}{2} \cdot \|V_\Sigma\|_2 \cdot \frac{N}{NumThreads}$$

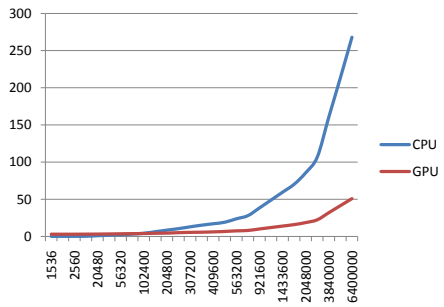


Figure 1. The speed of calculating a linear approximation for the input data (vectors) in milliseconds. Comparison of speed of calculation on the CPU and GPU.

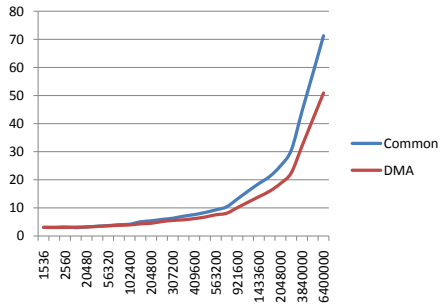


Figure 2. The speed of calculating a linear approximation for the input data (vectors) in milliseconds. Comparison of the effectiveness of implementation of calculation on the GPU using DMA access and common access to copy data to the global memory.

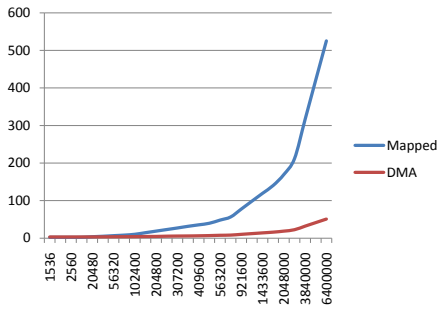


Figure 3. The speed of calculating a linear approximation for the input data (vectors) in milliseconds. Comparison of implementations using the mapping page-lock memory into the address space of the CUDA and approach using the DMA to copy data to the global memory.

VI. Conclusion

This article presented a parallel implementation of polynomial approximations on the GPU, which will significantly optimize the performance during calculation the large amounts of data. The implementation was compared with single thread CPU implementation, which is more suited for smaller data amount. It was shown that copying data from RAM, allocated as a page-lock memory, using the direct memory access (DMA), significantly accelerated the application performance in the result. In contrast, the use of mapping page-locked memory into address space of the CUDA in the implementation provided no improvement in application performance. This method is suitable for integrated GPU, which almost always produces a positive result due to the shared memory of CPU and GPU.

Acknowledgment

The thanks belongs to Professor Václav Skala for his substantive comments.

References

[1] NVIDIA Corporation, *CUDA ZONE*, http://developer.nvidia.com/, retrieved: December, 2011.

[2] NVIDIA Corporation, *NVIDIA CUDA C Programming Guide*, Version 4.0, 2011.

[3] G. Coombe, *An Introduction to Scattered Data Approximation*, October 31, 2006.

[4] K. Rektorys, *Survey of Applicable Mathematics II*, 7th ed. Praha, 2003.

[5] J. Sanders and E. Kandrot, *CUDA by example: an introduction to general-purpose GPU programming*, 1th ed. United States of America, 2011.

[6] M. Harris, *Optimizing CUDA*, SC, 2007.

[7] R. Farber, *CUDA, Supercomputing for the Masses*, May 14, 2009, http://drdobbs.com/high-performance-computing/217500110, retrieved: December, 2011.

 ISBN: 978-1-61208-176-2

Generating Context-aware Recommendations using Banking Data in a Mobile Recommender System

Daniel Gallego Vico, Gabriel Huecas and Joaquín Salvachúa Rodríguez
Departamento de Ingeniería de Sistemas Telemáticos
Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid
Avenida Complutense 30, 28040, Madrid, Spain
Email: {dgallego, gabriel, jsalvachua}@dit.upm.es

***Abstract*—The increasing adoption of smartphones by the society has created a new area of research in recommender systems. This new domain is based on using location and context-awareness to provide personalization. This paper describes a model to generate context-aware recommendations for mobile recommender systems using banking data in order to recommend places where the bank customers have previously spent their money. In this work we have used real data provided by a well know Spanish bank. The mobile prototype deployed in the bank Labs environment was evaluated in a survey among 100 users with good results regarding usefulness and effectiveness. The results also showed that test users had a high confidence in a recommender system based on real banking data.**

***Keywords*-Mobile Recommender; Context-aware; Banking data mining; User modeling; Customer segmentation**

I. INTRODUCTION

In recent years the mobile world has evolved extremely quickly not only in terms of adoption, but also in technology. The result of these advances is a high adaptive personalization of mobile applications. These new capacities provided by smartphones give rise to the possibility of building enhanced mobile commerce applications using all the user data we have at our disposal by utilizing their context sensors.

On the other hand, the eBusiness world has also advanced due to this new way of personalization. Good examples of this evolution are recommender systems. Traditional recommender systems usually are based on subjective data or personal scores provides by the users (e.g. Google Places). However, in recent years new platforms have based their recommendation on real purchases and therefore, the recommendations inspire more confidence (e.g. Amazon). This confidence in the results is always a key feature in any recommender system, but usually it is not easy to have such kind of data from real purchases. As a result, if we think in bank entities, we will probably agree that they are one of the best sources of trusted data in the world, as they have a huge amount of transactions from millions of users.

In this paper we present a mobile prototype based on using banking data to generate enhanced context-aware recommendations. This research project was carried out through the collaboration between our research group and one of the most important Spanish banks (its identity is not revealed in these lines in order to comply with bank's policies). This banking entity has provided us with more than 2.5 million credit card transactions made during the year 2010 and information about the 222,000 places and 34,000 anonymous customers' profiles related to the previous transactions.

The rest of the paper is organized as follows: the next section reviews related work. Section 3 describes the motivations behind this research. Section 4 presents the model used to generate the context-aware recommendations. Section 5 provides the results of our experimental work based on the prototype deployment and the survey carried out. After that, in Section 6 we discuss the results achieved. Finally, we conclude with a short summary and directions for future research.

II. RELATED WORK

A large amount of research and practical applications exist on mobile computing, recommender systems, context-awareness (e.g. [1] or [2]) or location based services, as well as any combination of the above areas. For instance, Kenteris et al. recently surveyed the field of mobile guides [3]. Ricci also discusses the goals of context-dependent recommendations and their importance in mobile recommender systems in his recent survey [4].

However, as Yujie and Licai stressed in [5] one of the most important challenges for context-aware recommender systems is the lack of public datasets available to conduct experimental evaluations on the methods developed.

On the other hand, it is important to note that usually all of the projects related to a banking data mining process in a bank entity are focused on generating mined knowledge useful for the bank workers, helping them to make decisions about customer segmentation or economic products, as we can see in [6] and [7]. Therefore, we cannot find research work in which the banking data is used to generate recommendations to the end users.

 ISBN: 978-1-61208-176-2

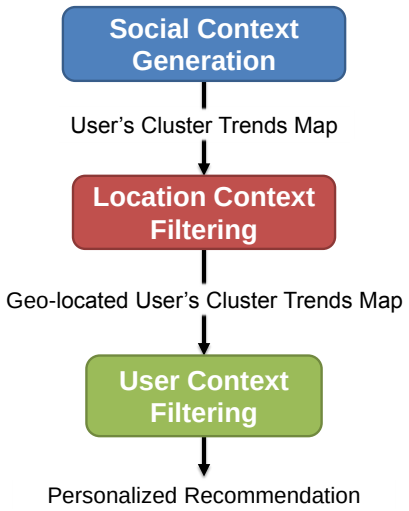


Figure 1. Adaptive recommendation process.

III. Motivation

All the important banks have millions of records in their databases plenty of rich information about customer purchases, client profiles or economic trends. Nevertheless, the vast majority of them frequently do not use correctly or underutilize these data to achieve a maximum benefit for their clients. Sometimes this is so because the privacy and security policies related to these data are complex to manage. In other cases, the challenge is related to a data mining scalability problem due to the huge data available to be processed.

Bearing this in mind, the emerging of recommender systems in this environment in response to these problems is a direct consequence. We have developed a method to generate context-aware recommendations for mobile recommender systems based on banking data. With this enhanced context-awareness our aim is to recommend *places*. A place is any entity where bank clients have paid with their credit cards (e.g. restaurants, stores, cinemas, supermarkets and so on).

Consequently, we have achieved a novelty application in the banking environment. This extra value provided to the end users is essential for our Spanish banking partner as it is an advantage in terms of market competition. It is also important to point out that using these banking data span across a wide domain range of recommendations categories, whereas most prior work tends to be more narrowly, often focused on a single store or a small set of products. Hence opportunities arise for cross-domain recommendations due to the richer context is possible to generate using banking data.

An additional main idea behind this research is the confidence on the recommendations generated. As Swearingen and Sinha [8] and Tintarev and Masthoff [9] said, one of the key goals of every recommender system is achieve the *trust* property to increase the users' confidence in the system recommendations. When we usually use a recommender system, we can think about several ways of falsifying or distorting the reality related to the items recommended. For example, the score of a restaurant recommendation from Google Places [10] is based on different user opinions. Thus the final recommendation is based on subjective evaluations of each user, and in some cases, the recommendation might not correspond to the reality. In conclusion, sometimes you might not trust recommendations because of the doubtful data origin. In our case we accomplish this goal because the system inspires confidence, as the data used for recommending are real data from the bank.

IV. Context-awareness Generation

As we mentioned in Section 2, there has been much research on the area of generating context-awareness and different definitions of the term *context* exist (e.g. [11], [12]). Therefore, we follow the definition proposed by Dey [13]: "*Context is any information that can be used to characterize the situation of an entity*". Specifically, the context dimensions in which our system is based on are: Social, Location and User context.

In the following sections we are going to present how we generate and use them to improve the recommendations, describing in detail the adaptive recommendation process summarized by Figure 1.

A. Social Context

The social context is generated by a data mining process over the banking data divided into three steps (Figure 2). These steps are not constricted to a real-time execution because all of them are carried out before the recommendations are requested by the user.

In the first step (User Profile Clustering), the system takes the banking client profiles provided to apply a clustering segmentation to them. These data have to be cleaned before the processing starts, so each record containing incorrect data (e.g. incorrect format, missing values, etc.) are ignored in the clustering process. It is very important to point out that only a restricted set of information was provided by our Spanish banking partner from its databases, being also previously anonymized in order to avoid privacy problems and to comply with the bank's policies. For this reason, the client profiles provided by the bank entity are represented by the following straightforward *n*-tuple:

$$< profileID, gender, age, averageExpensePerYear > \quad (1)$$

In order to reduce the complexity of the banking data mining process, we first apply a Canopy clustering process [14] on these data and then a K-means clustering process [15] over the canopies generated, achieving a set of clusters based on the similarity of the banking clients. We have

 ISBN: 978-1-61208-176-2

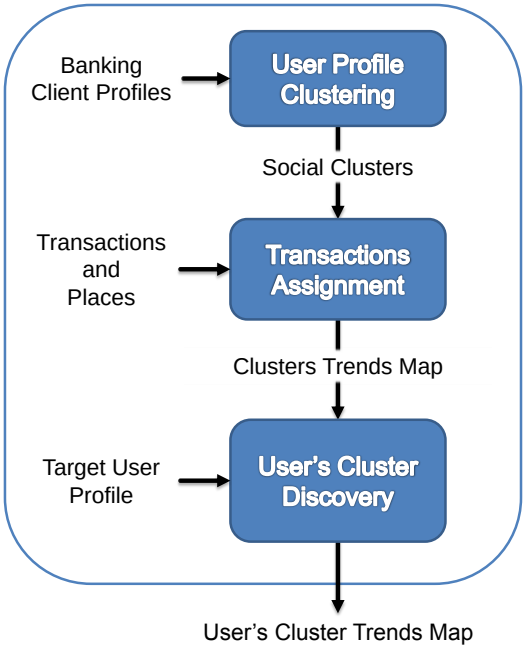


Figure 2. Social context generation process.

called this set of clusters *Social Clusters*, because they gather together banking clients with similar profiles, forming social groups where the consumption model or tastes are related.

In the Transactions Assignment step, the system first assigns the credit card transactions to the corresponding cluster, considering that there is a unequivocal relationship between a transaction and a client (given by the *profileID* element that indicated who made that credit card transaction). Every bank transaction is represented by the following *n*-tuple:

$$< profileID, placeID, paymentAmount, time, date > \quad (2)$$

After that, all the transactions are assigned to the social clusters and then, a second process identifies the places where the transactions were made. The places are represented by the following *n*-tuple:

$$< placeID, category, name, address, latitude, longitude > \quad (3)$$

With this second process, we create a map of places where the relationships among places and clusters are shown, noticing in this way the consumption trends of every cluster.

Finally, the User's Cluster Discovery process is activated when the user enters the first time to the mobile application. The system checks the information profile extracted from the user's banking account (a *n*-tuple like the one show in 1) in order to assign her to any of the existing social clusters. This is carried out by calculating the distance among the point

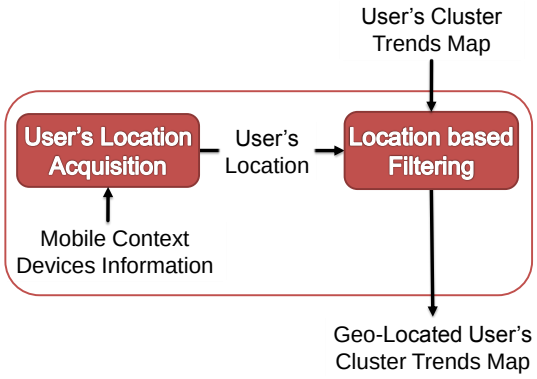


Figure 3. Location context filtering process.

that represents the user profile in the space defined and the centroids from every social cluster. A centroid is a virtual point corresponding to the average of all the real points in the cluster. That is, its coordinates are the arithmetic mean for each dimension separately over all the points in the cluster. Hence, the cluster with the centroid at a minor distance from the user profile point representation is the social cluster assigned to the user.

After these steps we know the social context of the user because now she has been assigned to one of the social clusters generated. Every cluster has a common consumption model represented by the Clusters Trends Map and thus, we know which places are candidates to be recommended to her.

For instance, if a user belongs to the social cluster of 50 year-old women with an average expenditure of EUR 10,000 per year in credit card purchases, the set of possible places to recommend is made up of the places in which people in this category have paid with their credit cards in the year 2010 (as the data provided by our bank partner for this research correspond to that year).

B. Location Context

As [11] said, location is currently one of the most important context parameters. Accordingly, after obtaining the social context of the user based on the banking information, the recommendation can be made more accurate by adding the location context dimension. Most of the time, end users are looking for places recommendations in their immediate locality (e.g. good restaurants nearby). The use of mobile context device information as an input for the recommender system allows us to personalize recommendations based on the user's location.

Different mobile context devices are involved in the acquisition of the user's location. If the user's device is GPS-capable, the geo-location will be more accurate. If not, a less accurate but usually valid location can be obtained using network-based positioning technologies or Geo IP capabilities.

 ISBN: 978-1-61208-176-2

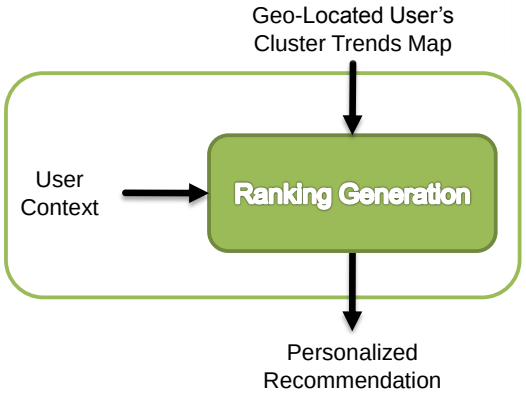


Figure 4. User context filtering process.

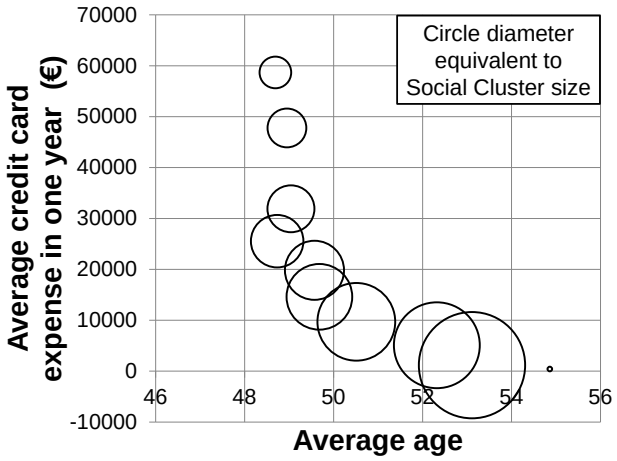


Figure 5. Social Clusters distribution by average expense, age and size.

Once the system is aware of the user's location, it is applied as a new input to filter the user's cluster trends map, obtaining a geo-located user's cluster trends map (Figure 3).

C. User Context

The final process to achieve the personalized recommendation takes into account the user context. This context is based on a set of parameters (e.g. current time or current activity of the user inferred from sensor data) and an input preference given by the user to know the place category (one of the elements of the *n*-tuple 3) she wants to be recommended (e.g. restaurant, supermarket, cinema, etc.).

For instance, if the user wants a restaurant recommendation (category input), the mobile application could also use the current time information (e.g. lunch time) to filter the geo-located user's cluster trends map (Figure 4) considering only those restaurants that fit with her current activity (e.g. walking). Following with the example, the user would see only a ranking of the closest restaurants at walking distance to her location that the banking clients belonging to her social cluster has visited the most at lunch time. That ranking would be generated by ordering those restaurants attending to the number of customers that have previously visited everyone.

V. Evaluation and Results

To evaluate the system, a prototype was developed and deployed in a real environment that belongs to our bank partner called *Labs*. The primary aim of Labs is to allow the deployment of new researches and development projects created in the bank in order to be able to collect feedback from a set of bank clients registered in this environment.

Using this platform we evaluated first the social clusters achieved after applying the previous processes to real banking data. Then, we set up an online survey with two scenarios using a mobile prototype developed in Android in order to evaluate the user acceptance.

A. Social Clusters

The banking data provided by our bank partner to create the social clusters consisted of more than 2.5 million credit card transactions made during the previous year, providing information on 222,000 places and 34,000 anonymous customers' profiles from customers between 48 and 55 years old. All these data were provided by the bank following the n-tuples (1), (2) and (3).

The Figure 5 illustrates the social clusters emerged after the clustering process over the banking data. As we can see, the average credit card expense in one year, the average age and the size of every social cluster (given by the diameter of the circles) is shown.

B. User Acceptance

We have analyzed the feedback provided by 100 bank customers registered in the Labs platform where the system is deployed. This evaluation was carried out using an on line questionnaire based on two scenarios. The first one was focused on restaurant recommendations and the second one on supermarket recommendations. Both scenarios show a simple case in which after a user request, she receives a recommendation compose by several places corresponding to the previous categories. Figure 6 depicts a screenshot for a lunch recommendation provided by the Android mobile prototype taking into account that the location is provided by the mobile context sensors and the user context information is previously provided by the user.

Therefore, after a brief experience with the application in those scenarios the test users were asked to judge several statements related to some properties using a 5-point scale, where 1 mean *"totally disagree"* and 5 *"totally agree"*. The statements were like this: *"The application is [property evaluated]"*. Additionally, users had free text inputs fields to

 ISBN: 978-1-61208-176-2

Figure 6. Mobile application interface for a lunch recommendation.

make comments and annotations. The results are illustrated with average values in Figure 7.

VI. Discussion

First of all, if we think about the distribution of customers along the social clusters (Figure 5), the results confirm the intuition: the clusters with less people are the clusters who spent more money in credit card transactions because high economic class people are more infrequent than medium and low economic class people. While the bigger size clusters are those who spend less money and also, are composed by older people that are less used to pay with credit card than younger people.

In regard to the results related to the user acceptance, they reveal a very positive attitude towards the mobile recommender system shown in the Figure 6, as long as it has average high scores in all the properties analyzed.

On the other hand and attending to the way we manage the explanations in our recommender system, it achieves some of the most important criteria recently set out by Tintarev and Masthoff in [9]. Specifically the "transparency" (i.e. explain how the system works) is achieved due to the explanation provided for the places recommended, as the application informs the user about how the recommendations have been generated considering the purchases of other bank customers like her. The "trust" (i.e. increase users' confidence in the system), "effectiveness" (i.e. help users

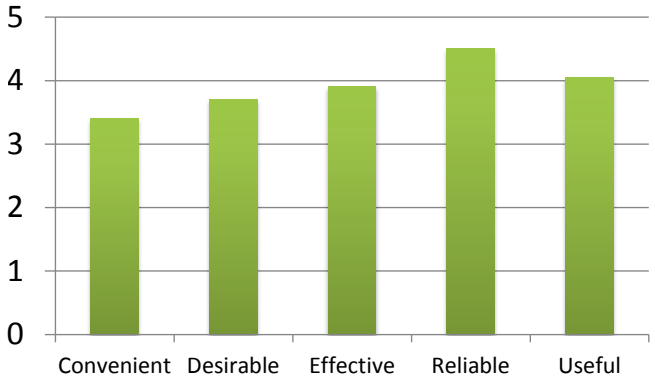


Figure 7. User acceptance survey results.

make good decisions) and "satisfaction" (i.e. increase the ease of use or enjoyment) criteria are achieved if we take into account the high values of the "reliable", "effective" and "useful" properties respectively evaluated in the survey (Figure 7).

The statistical outcome is also supported by the comments wrote by the test users during the survey process. For example: *"Having a recommender system in my smartphone available anywhere, anytime for searching any kind of place is very usefull."*, *"I think that the most important feature for me is the confidence in the results as they come from real data of my own bank"* or *"I will really appreciate to have such a kind of application in my mobile phone in my daily life."*.

However, some of them remarked the privacy issues related to deploy this system into a real commercial exploitation, as they did not want their personal data in danger. Although the system is right now deployed in the bank Labs environment (a close secure environment), this point is an important issue that has to be studied if the system is deployed in a future outside of it, attending to the facts pointed out by Ohm in [16], where he showed that in some cases anonymization is not enough to preserve privacy promises.

VII. Conclusion and Future Work

In this paper, we have presented a method of generating context-aware recommendations using banking data in mobile recommender systems. As we have shown, using this kind of data based on real people actions and banking history, allows us to increase the confidence in the personalized recommendations generated, because there is no subjective data used in the recommendation process. This feature of our system provides an essential advantage compared to other recommender systems which have the aforementioned problem of being based on non-reliable data.

Current and future work includes the evolution of the current prototype into a complete mobile application. This would allow us to evaluate the system with users really

 ISBN: 978-1-61208-176-2

interacting with a mobile device in realistic scenarios related to the users' daily life. Of course, the long run aim of our banking partner is to launch a real product for commercial exploitation based on our system. Thus it will be necessary to study the privacy problems related to that real deployment.

On the other hand, we want to analyze the impact of using proactive techniques in our recommendation process. Proactivity means that the system pushes recommendations to the user when the current situation seems appropriate without being needed a explicit user request. Therefore, we are working now in a model to achieve proactivity in mobile context-aware recommender systems ([17] and [18]) that could be integrated into the context-generation model presented in this paper.

Another open issue that could be studied in relation with enhancing the recommendation process is the one described in [19], in order to create multiple personalities in the system that would have their own personalized profile, like a "what kind of customer do you want to be today?" feature. As a result, a user of the system could have several profiles with different social clusters associated. This is an interesting feature if we bear in mind that sometimes people pay with their credits cards to buy gifts or services for friends or family that usually does not have the same tastes.

ACKNOWLEDGMENT

The authors would like to thank the bank *Labs* group for providing their banking data and their expertise on the banking domain without which this research could not have been possible.

REFERENCES

[1] M. Baldauf, S. Dustdar, and F. Rosenberg, "A survey on context-aware systems," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 2, pp. 263–277, June 2007. [Online]. Available: http://portal.acm.org/citation.cfm?id=1356236.1356243

[2] G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," in *Proceedings of the 2008 ACM conference on Recommender systems*, ser. RecSys '08. New York, NY, USA: ACM, 2008, pp. 335–336. [Online]. Available: http://doi.acm.org/10.1145/1454008.1454068

[3] M. Kenteris, D. Gavalas, and D. Economou, "Electronic mobile guides: a survey," *Personal Ubiquitous Comput.*, vol. 15, pp. 97–111, January 2011. [Online]. Available: http://dx.doi.org/10.1007/s00779-010-0295-7

[4] F. Ricci, "Mobile recommender systems," *International Journal of Information Technology and Tourism*, 2011.

[5] Z. Yujie and W. Licai, "Some challenges for context-aware recommender systems," in *Computer Science and Education (ICCSE), 2010 5th International Conference on*, aug. 2010, pp. 362 –365.

[6] P. Ataee, "Mining the (data) bank," *Potentials, IEEE*, vol. 24, no. 4, pp. 40 – 42, 2005.

[7] S. Ren, Q. Sun, and Y. Shi, "Customer segmentation of bank based on data warehouse and data mining," in *Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on*, 2010, pp. 349 –353.

[8] K. Swearingen and R. Sinha, "Interaction design for recommender systems," in *In Designing Interactive Systems 2002. ACM*. Press, 2002.

[9] N. Tintarev and J. Masthoff, "Designing and evaluating explanations for recommender systems," in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Springer US, 2011, pp. 479–510.

[10] Google, "Places," 2011. [Online]. Available: http://www.google.com/hotpot

[11] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, ser. HUC '99. London, UK: Springer-Verlag, 1999, pp. 304–307. [Online]. Available: http://portal.acm.org/citation.cfm?id=647985.743843

[12] M. Bazire and P. Brézillon, "Understanding context before using it." in *CONTEXT'05*, 2005, pp. 29–40.

[13] A. K. Dey, "Understanding and using context," *Personal Ubiquitous Comput.*, vol. 5, pp. 4–7, January 2001. [Online]. Available: http://dx.doi.org/10.1007/s007790170019

[14] A. McCallum, K. Nigam, and L. H. Ungar, "Efficient clustering of high-dimensional data sets with application to reference matching," in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, ser. KDD '00. New York, NY, USA: ACM, 2000, pp. 169–178. [Online]. Available: http://doi.acm.org/10.1145/347090.347123

[15] T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 881 –892, Jul. 2002.

[16] P. Ohm, "Broken promises of privacy: Responding to the surprising failure of anonymization," *Social Science Research Network*, vol. 57, pp. 1–64, 2009.

[17] W. Woerndl, J. Huebner, R. Bader, and D. Gallego-Vico, "A model for proactivity in mobile, context-aware recommender systems," in *the 5th ACM International Conference on Recommender Systems*, October 2011.

[18] D. Gallego-Vico, W. Woerndl, and R. Bader, "A study on proactive delivery of restaurant recommendations for android smartphones," in *the International Workshop on Personalization in Mobile Applications (PeMA) at 5th ACM International Conference on Recommender Systems*, October 2011.

[19] L. F. Cranor, "'i didn't buy it for myself' privacy and ecommerce personalization," in *Proceedings of the 2003 ACM workshop on Privacy in the electronic society*, ser. WPES '03. New York, NY, USA: ACM, 2003, pp. 111–117. [Online]. Available: http://doi.acm.org/10.1145/1005140.1005158

 ISBN: 978-1-61208-176-2

Web Personalization

Implications and Challenges

Ahmad Kardan, Amirhossein Roshanzamir
Department of Computer Engineering and IT
Amirkabir University of Technology
Tehran, Iran
aakardan@aut.ac.ir, amrhssn@aut.ac.ir

***Abstract* — Companies are under the pressure to provide tailor-made products or services that match customers' preferences better. Personalization from web mining is a significant tool to accommodate this trend by extracting patterns of customer's preferences and connecting them directly to production line and supply chain. However, with exception of companies like Amazon, Dell, Toyota and most of Airlines which have solid strategy and large appeal, it is challenging to develop a concrete and cost-effective approach in personalization for many companies. This paper studies dimensions of personalization and addresses business implications and challenges of web personalization. It will further suggest a novel approach for developing web personalization in small and medium sized enterprises by utilizing Frequent Flyer Program of Airlines Industry in order to justify and guide investment.**

Keywords-web personalization; build-to-order; customer relationship management; recommender system; frequent flyer program.

I. INTRODUCTION

Web Personalization is simply defined as the process of gathering and storing data about visitor's interactions and navigation in a website in order to assemble and deliver the tailor-made web experience to a particular user. The delivery can range from making the web site more appealing to anticipating the needs of a user and providing customized and relevant information. The experience can be as simple as browsing a web site up to trading stocks or purchasing a computer. Effective personalization can be achieved by three steps including identifying, retrieving and assembling. The first step starts by collecting all available data. In fact, the user's data is divided into two categories i.e. personal data such as age, gender and demographics as well as behavioral data such as usage, click stream and time [1]. In the next step, the web site forms the visitor's profile and utilizes intelligent algorithms to analyze and mine data in order to extract statistics, and discover correlations between web pages and user's preferences. In the final step, the web site will deliver the right information and/or produce customized products and services to meet each user's requirement or assemble the most preferred page to be displayed to his or her preferences. This paper reviews dimensions of personalization and then addresses business implications and challenges of web personalization. It will further suggest a novel approach for developing web personalization in small and medium sized enterprises by explaining a mobile portal as a case study. A Frequent Flyer Program of Airlines Industry is used to justify and guide investment in this approach.

II. DIMESIONS OF PERSONALIZATION

In practical applications, we can suggest a model to divide personalization technologies into two dimensions namely horizontal and vertical. Horizontal dimension enables a company to adapt attributes and appearance of products or services according to customer's flavor and taste, whereas vertical dimension enables a company to interact with customers in order to customize the configuration, performance and quality of products or services according to customer's request .

A. Solid

This is a kind of personalization which a user has no control to change or modify the product, nor is there any interaction between website and user. The user or customer simply selects the product and then use it. Digital music and video store and most of retail online sites which sell specific and predetermined products are examples of this type .

B. Superficial

This is a kind of personalization in which user has only limited control on appearance and presentation of the products. Many portal sites, such as Yahoo, Google and MSN allow users to personalize the page with selected news, local weather forecast, and other features. For instance, web news portal may provide news and articles about education and business school ranking to customers who are avid MBA fun and sport news for sportsmen. Digital Greeting Cards stores are another example where the user can select the type and structure of greeting card and can further customize it by individual messages.

C. Evident

This is a kind of personalization which user has no control to modify the product however, he or she can communicate with supplier and inform his or her specific interests while receiving recommendations and reviewing buying patterns of other customers of similar preferences. The customer can further place order for his or her preferred type or version or predetermined price. Amazon is an example of this kind in which users can receive

 ISBN: 978-1-61208-176-2

recommendations and review comments of other customers. They can also request for digital version of a book and once the total requests reach a certain number, Amazon will fulfill this request. It is actually, manifestation of one of Amazon's strategies titled obsess over customers, to start with the customer and then go backwards to develop product and service solutions [2]. eBay can be considered as another example of evident personalization in which consumers and businesses engage in buying and selling a variety of goods and services worldwide. eBay online portal facilitate these ventures by providing Auction-style listing in which buyers and sellers can set and adjust their specific requirements and predetermined price.

D. Collaborative

This is the ultimate level of personalization in which user has control in designing and building the product according to his or her preferences. User usually selects the modules that are literally building blocks in order to customize a product and then assembles various combinations of modules. In electronics examples of modules would include processor, mother boards, memory, disk drives and software. Dell Computers has created a unique process within Industry and pioneered the build-to-order computer business. The process was long and required a great change in the thinking of many firms and many people within Dell and it has taken 20 years to get where it is [3]. Toyota, for example, introduced Buyatoyota.com as the major step in order to integrate its unique Just In Time (JIT) system with personalization and customer's specific requirements. This site guides the customer through the steps of selecting a model, viewing and searching options, assembling the features and specifications, choosing the color and accessories, and finally, locating a local dealer that can provide that choice of car in order to get a quotation and arrange finance. Other car manufacturers may start to use the same approach in building their cars in near future.

Vertical Personalization \ Horizontal Personalization	
Evident e.g. Amazon, eBay	Collaborative e.g. Dell, Toyota
Solid e.g. Online Music	Superficial e.g. E-card, E-News

Fig 1. Dimensions of Personalization

III. PERSONALIZATION-BROADER STRATEGY

In fact, personalization must be an integral part of a broader strategy and connect to Customer relationship Management (CRM) and Recommender Systems. CRM is viewed as a strategy to attract, grow and retain customers. Personalization is an approach that can aid in bringing, staying and returning visitors / customers to a website. Since the very nature of the web tools encourage interactivity between people and organizations, the topics of personalization and CRM are, therefore, complementary to each other. They both have the ability of providing the right information or content (e.g. products, services and data) to the customer at the right time and right place [4].

The explosive growth of e-commerce and online environments has overwhelmed users by countless options to consider. They simply neither have time nor knowledge to personally consider and evaluate these options. Recommender systems represent a prominent class of personalization applications that aim to support online users in deciding which products or services meet their requirements. The advanced version of a recommender system, for example the one implemented in Amazon Inc. , would add information about other complementary products (cross selling), in the form other customers that had bought X, also had bought Y and Z. Today, recommender systems have become one of the most powerful and popular tools in electronic commerce, since, they allow to transfer users into customers, increase the cross selling and build loyalty [5].

IV. FREQUENT-FLYER PROGRAM

A Frequent Flyer Program (FFP),which is a loyalty program offered by many airlines, was first created by Texas International Airlines in 1979. FFP can be considered as an interesting manifestation of evident personalization which integrates personalization with CRM and recommender system. Passengers can typically enroll in FFP of an airline and accumulate miles based on the distance flown of that airline or its partner. Miles accumulated allow members to redeem tickets, upgrade service class, or obtain free or discounted car rentals, hotel stays, merchandise, or other products and services through partners. The personalization features include but not limited to the possibility of selecting seats through online portal and requesting for special meal. Members can manage their account online by buying tickets and checking online while receiving personalized news and special offers based on their preferences and destination flown. In recent empirical study on FFP, Drèze and Nunes argue that successful redemption of miles fosters reengagement of passengers and motivates them to flying more frequently. Therefore, loyalty programs that offer people multiple redemption opportunities must balance the attractiveness of a reward with an appropriate level of difficulty in attaining success [6]. This research further shows that loyalty can be better accomplished when the reward is not too difficult to achieve, rather it should be inspiring and challenging to cash out and when someone does, he or she feels successful.

 ISBN: 978-1-61208-176-2

V. IMPLICATIONS AND CHALLENGES

Personalization has gone through different development phases since early 2000. It initially started as a tool to attract and retain visitors by giving them chance to explore more of the site. Advertising and promoting products and services, nevertheless, were part of this phase as well. The next phase attempted to increase turnover of customer's spending by offering more expensive or similar products. Today, personalization is increasingly used as a means to speed up the delivery of the right information to a visitor in order to customize products and services for meeting and exceeding his or her requirements. Those companies that are systematically gathering information about their customers, product attributes, purchase contexts and integrate it with behavioral segmentation such as demographics, attitudes and buying patterns can make more sophisticated offers that identify customers who are most likely to defect [7]. This personalization strategy ultimately increases number of regular customers and amount of each transaction and has made personalization as a required and expected feature of an e-business. For example, Timberland Boot Studio by allowing customer to select different leathers and colors gets three times hits on its customized boots [8].

Without challenging the enormous potential and contribution of personalization technology, the question still remains whether personalization really works? The smart answer is it depends. On one hand the benefits could be significant not only for the website visitor (being offered more interesting, useful and relevant web experience) but also for the provider (allowing one to one relationships and mass customization and improving website performance). On the other hand, personalization requires extensive and precise data that are neither easily obtainable, nor can be mined and analyzed efficiently. As such in many cases the output does not seem so successful in understanding and satisfying user needs and goals.

First and foremost, the ethical dimension of personalization need to be addressed, since online user's navigations are recorded for building and updating user profiles and this can put privacy of users in jeopardy. At the same time users are becoming more vigilant and have higher expectation. They are not so happy with idea of being stereotyped without their consent. Users also expect to be treated equal and have enough control and choices. The cost of technology initiated from intelligent software and storing hardware as well as the time spent are also critical factors which must be justified in the long run. Schneider and Bowen [9] proposed to explain that customer satisfaction of the service originates with the handling of three basic customer needs: security, justice and self-esteem. Building on their three needs conceptualization, we suggest six major implications and challenges of personalization as follows :

A. Security and Privacy

There are implications for user's privacy and security of information, since, much of personalization entails intensive collection and use of personal information. The terms "privacy" and "security" are often used interchangeably, but there is an important distinction between these two. Security refers to the ability of user or site to protect information against unauthorized third parties by preventing them to access, use or modify information whereas privacy is the quality of being secluded from the presence or view of others [10]. It further refers to the ability and the right of the users to decide and control what can happen to their data and profile. Given the importance of data acquisition to personalization approach, it is crucial that sites identify privacy preferences of users and the relationship to their satisfaction with the site. This includes the users' level of acceptance in how the data is acquired, whether the benefits of the approach outweigh the privacy risks, and whether the site will disclose the information to third parties [11]. As such users, at least, must explicitly feel connected to information in order to start benefiting from the service/features. Once users see benefits, they might be willing to surrender additional information and be more transparent, provided they know what is going to be done with it. In addition, the site must take all measures against the factors that could be outside of the knowledge or understanding of the user, such as the sufficiency of security mechanisms to protect any data provided. The user may often be unaware of what data is being given away and how they are stored and secured in the case of passive collection of information. Therefore, the site must encrypt passwords and sensitive data of users and evaluate an external test to ensure about the security and protection of data. A commonly recommended practice is to declare a privacy statement (or disclosure statement) which describes exactly what kind of data are gathered from users and then declare the policies about methods of using and sharing it .

B. Fair dealing and Integrity

These are critical issues in dealing with customers and visitors of a web-site. In fact, all visitors expect to be treated equally in terms of information, prices and services provided. When it comes to personalization, there are of course many occasions that a company is obliged to act discriminately based on the time, efforts, and money invested by customers as well as level of loyalty and pervious transactions. In such cases, the reason for differentiation or privileged treatment must be publicly announced to address user's expectation for equality and avoid misunderstandings.

A good example of implementing this discriminatory practice is in Airline Industry which employs different prices based on booking time and publicly announces it. Passenger might have paid different fair tickets but upon boarding they would receive similar treatment. Today, fair dealing, integrity and justice is becoming critical issues and few other factors such as keeping commitment and flexibility in dealing with unusual requests are also emerged accordingly. Personalization has the capacity to reinforce all these based on the nature of business and

 ISBN: 978-1-61208-176-2

dimensions of personalization. It can also meet reasonable yet abnormal requests by keeping records and commitments .

C. Self-esteem and Sense of worth

Personalization can provide a unique possibility to maintain and enhance self-esteem and sense of worth for visitors by providing a user-friendly platform in which people feel in control, important and comfortable while having enough choices. Maslow claimed that the need for self-esteem can be met through mastery or achievement in a given field or through gaining respect or recognition from others [12]. As users become more vigilant, they might find personalization more pleasant and appealing, if they could exercise more control over it. We can imagine a scenario in which a online bookseller asks a visitor, "Would like us to add this title to our growing knowledge of your interests to lead and direct our future recommendations?" The customer can select "yes" and enjoy receiving recommendation, if he or she is an avid reader fun of the same subject. Alternatively, he or she can say "no" and spare time of receiving and reading these recommendations [13]. Likewise, a website which has personalization capabilities to discover and analyze the patterns in customer's navigations can fulfill their need by saying "We have noticed you frequently check football news, would you like us to update this news on your home page?" In both these scenarios, regardless of visitors response, we are trying to view them as unique people by respecting their interests and self-esteem while giving them some control over displayed contents. Even, a simple greeting message by indicating the name of visitor can enhance customer's feeling of self worth. This approach which is simple yet powerful can address privacy requirements of vigilant users, since, personalization is done after receiving their consent .

D. Cost

Successful companies like Google, Amazon, Dell and most of airlines spent millions on their portal web-site in order to personalize their products and services for their customers. In fact, personalizing online offerings is considered less costly than customizing physical products because of the “digital” nature of information goods. That is, with advanced information technology, online pages can be manipulated easily to suit individual customers’ needs. But, it all depends on the number of visitors and customers as well as the nature of business. Each individual need to have a profile with details of their preferences and every time he or she goes online and visits the site, the profile needs to be updated. Moreover, intelligent algorithms must be applied in order to extract the most suitable and customized products or services to be offered. There will be also interrelated links between preferences of users who have similar purchasing behavior. Although the price of using new technology is being reduced day by day, however, developing new algorithms and using thousands of servers to address each and every customer's need is time consuming and costly. It was found that operating a personalized web site can cost more than four times than operating a “comparative dynamic site” and most sites that deployed personalization have not realized adequate returns on their investments [14]. As such, there must always be a balance between cost spent and potential income which is likely generated by economy of scale and addressing core business strategy. Google for example has almost billion visitors every day using its search engine [15]; therefore, they can easily afford heavy investment on technologies such as collaborative filtering, data mining, and click-stream analysis in order to customize their offerings at the individual level. Amazon, nevertheless, has thousand customers too and personalization is an important part of their sales strategy .

E. Timing

Web personalization enables online websites to customize their contents by capturing real time preferences of individual visitors through web mining techniques. The next step is to adapt the website content in order to meet individual's specific requirements. Yet, there is a trade-off between quality of recommendation and probability of accepting a given recommendation i.e. although the content of web-site will improve during the course of session to meet preference of the visitors, however, the probability of using and enjoying the fresh contents diminishes over the course of session [16]. These effects suggest that online portals have limited time to capture and mine visitors data in order to customize the most tailor-made contents or products. In fact, consumers prefer early presentation that eases their selection process, whereas adaptive systems can make better personalized content if they are allowed to collect more consumers' clicks over time [16].

Therefore, personalization needs to be efficient enough in order to keep the balance between time spent by user and to the extent to which his or her online behavior can be symbolized .

F. Agility

With advances in tracking and database technologies, companies can better understand and evaluate their customer's requirements. However, they need to build up certain capacity in order to rapidly translate this understanding into appealing products and service. Agility or nimbleness is the capability to swiftly adapt to changes and can be achieved in three distinct ways including operational, portfolio and strategic [17]. Operational agility is illustrated as the success of Toyota and Dell through integrating supply chain and directly linking customer to production line. While sharing real-time market data that is detailed and reliable, Dell and buyatoyota.com only assemble the product after receiving the order (build-to-order) and this strategy increases visibility to the demand and flow of goods. The primary advantage is sensitivity to changes in customer demand and possibility of mass-customization.

 ISBN: 978-1-61208-176-2

Toyota, for example, outsources about 70% of components and that is why Toyota Production System (TPS) requires a serious investment in building an agile network of highly capable suppliers of different components that must be truly integrated into supply chain [18]. In this scenario, the user goes through the process of designing a car by selecting features and components online. Once the order is registered and the car dealer and type of finance is arranged, the delivery of parts and components takes place a few times a day from different suppliers to Toyota factory and the car is specifically and solely assembled as per online order of the user .

Although a company's growth is dependent on finding and retaining customers, however, its success is far beyond what is on the web page and depends on internal operations (the back end) and its relationship with suppliers and other business partners [19]. The true power of such operational agility requires solid information technology infrastructure and is based on ingenuity and electronic supply chain management.

VI. GIVE UP OR BUILD UP

Today, companies can extract valuable information by exploiting information hidden in their web site created as a result of visitors interaction and browsing. Facing increasingly sophisticated customers, companies are under the pressure to provide tailor-made products or services that better match customers' preferences .

The major challenge is cost pressure and justification of investment by economy of scale. Except giant companies like Amazon, Facebook, Dell and most of Airlines, which have large appeal and thousands and even millions of customers, it is challenging to seek cost-effective approaches in personalization for small and medium sized enterprises. The appropriate solution for these companies could be to start their portal on Solid personalization basis and gradually develop it horizontally or vertically to Superficial and then Evident personalization based on the nature of business and improvement of the business model .

VII. WEB PERSONALIZATION METHODOLOGY

Commonly used to enhance customer service and e-commerce sales, personalization is sometimes referred to as one-to-one marketing, because the enterprise's web page is tailored to specifically target each individual consumer. The main purpose of every business is to make money either by selling products or services to new customers or selling more to existing ones [21]. Miller [20], in his book, argues that the average value of customers is 8 to 10 times their initial purchase depending on the research he and his colleagues have reviewed. He further argues that the cost to attract a new customer is 5 to 6 times more than your cost to save an existing customer.

As small and medium enterprises have serious challenge to create traffic on their site and increase the number of visitors, the best strategy would be to focus on existing customers by encouraging them to repeat the orders on similar or different products and services and/or cross selling. Frequent Flyer Program is a gifted tool which can be duplicated and applied in any online shop in order to grant points to the customer based on the value of their purchase. These points can be rewarded later in terms of free offering or discount scheme as incentives for repeated orders. By exploiting such a policy, a company can increase market share and improve profitability. Saving time and reducing costs of sales and marketing as well as economic use of resources are other indirect benefits of what can be called point plus program. All these will ultimately justify investment in web personalization. Business owner, Rosalind Resnick also insists on rewarding program by saying "It's nice to know that every dollar we spend to grow our business can also be a point or air mile we can use to celebrate our independence" [22].

In order to increase customer satisfaction and the likelihood of repeat visits, there must be a reason like membership. When people are members rather than shoppers, they feel connected and privileged. Therefore, the first step in formulating a solid personalization model is to encourage users to become members in order access special contents and/or perform special functions on the site while enjoying superficial personalization. Once a closer relationship with the users are developed, the company can gradually address the other challenges of personalization as indicated in Fig 2 and move to evident level and build interactive relationship to meet and exceed user's requirements. The collaborative requires ingenious infrastructure and agile supply chain, which certainly goes beyond information technology.

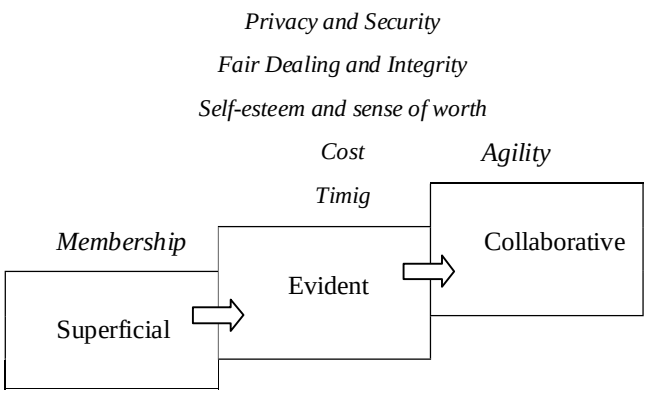


Fig 2. Web Personalization model

A mobile portal, which aggregates and provides contents and services for mobile users, is a good example for explaining how the above model works. These portals are specifically designed for m-commerce with short menu of popular topics like news, sports, email, travel information and finance as well as very few graphics. They recently

 ISBN: 978-1-61208-176-2

started to provide downloads, health, dating and job information as well [23]. In fact, a mobile portal provides an intimate one-to-one experience for a user who is visiting the portal for specific purpose like checking sports scores or looking to buy specific item in nearby shops [24]. Users frequently can become a member by paying a monthly membership to access basic information. This is superficial personalization in the above model. As users started to seek specific services and information, the portal need to address 5 challenges of web personalization in the same order to meet user's requirements for evident personalization. First and foremost, privacy and security are critical, since mobile systems tie highly personal information to location and contacts. We normally do not share our mobile device with anybody; therefore, anything that we do with our mobile is traceable to us. This may be helpful for marketing; yet, mobile advertisement is nowadays pursued with caution while getting the customer's consent. This can be done by offering long-distance in exchange for viewing the ads offered by mobile portals. Furthermore, mobile portals can customize advertising message for specific group of customers based on their location while knowing their preferences and buying habits. The integrity, fair dealing and self growth should be the essential cornerstones of these offers to convince for quick decision and it must be support by loyalty program to reward customers with points. The cost and timing are also hot topics for mobile portals. In addition to basic monthly fees, most mobile portals charge per-service fee for premium and customized contents such as download and weather forecast. The nature of m-commerce, in which users have little time to navigate the contents and wait for page to load, also demands that the content to be produced in much shorter time. We can imagine in the not so far future, Toyota and few other car manufacturers allow their customers to assemble and buy their favorite car through mobile portals.

VIII. CONCLUSION

Personalization from web mining has been received lots of interests in business as a gifted tool to improve sales and retain customers, since, it can increases customer‘s satisfaction by providing them with tailor-made products and services. Web Personalization can also help the company to implement build-to-order policy by connecting customer’s requirements and preferences directly to production line and supply chain. As a result the company can benefit from cost saving and efficient utilization of resources. Therefore, collaborative personalization can be considered as a vision of any small and medium sized enterprise to incorporate its online portal with built-to-order strategy for sales and marketing. This of course a long term plan which must be supported by integrating personalization into loyalty program in order to reward customers with points for a wide range of daily interactions with the company.

The reward scheme should be smartly designed to provide users with both mental and financial benefits in order to encourage them for repeated order and/or cross selling. The power and benefits of these two can create a solid momentum to justify and guide investment in personalization.

REFERENCES

[1] Kardan, A., Fani, M.R. and Mohammadian, N. (2011), Purposing an Architecture for Learner Modeling base on Web Usage Mining in e-Learning Environment, The 5th Data Mining Conference Dec. 14 - 15, 2011 ; Tehran, Iran

[2] Treanor, T. (2010), Amazon : Love Them ? Hate them ? Let's Follow the Money, June 2,2010 Springer Science + Business Media, LLC 2010, pp. 124-125

[3] Kumar, S. and Craig, S. (2007), Dell, Inc.'s closed loop supply chain for computer assembly plants,Information Knowledge Systems Management; 2007, Vol. 6 Issue 3, pp. 197-214

[4] Jackson, T.W. (2007), Personalization and CRM,Journal of Database Marketing & Customer Strategy Management (2007) 15, pp. 24–36

[5] Velaquez, J. D. and Palade, V. (2007), Building a knowledge base for implementing a web-based computerized recommendation system, International Journal on Artificial Intelligence Tools,Oct2007, Vol. 16 Issue 5, pp. 798

[6] Drèze, X. and Nunes, J. (2011), Recurring Goals and Learning: The Impact of Successful Reward Attainment on Purchase Behavior, Journal of Marketing Research (JMR), Apr2011, Vol. 48 Issue 2, pp. 268-281

Thomas H. Davenport, Leandro Dalle Mule, and John Lucker

[7] Davenport, T.H., Mule, L.D., and Lucker, J. (2011), Know What Your Customers Want Before They Do, Harvard Business Review, Dec2011, Vol. 89 Issue 12, pp. 84-92

[8] Turban, E. and Voloninio, L. (2010), Information Technology for Management, John Wiley & Sons, Inc. ;7 th Edition

[9] Schneider, B. and Bowen, D. (1999), Lessons in customer service, Understanding Customer delight and outrage Fall 1999, SloanSelection Winter 2011 pp. 4

[10] Becker, M. and Arnold, J. (2010), Mobile Marketing for Dummies, Wiely Publishing Inc.; 1st Edition

[11] Getek, R.C. (2010), A usability model for web-based personalization based on privacy and security, PhD dissertation ; University of Maryland, Baltimore County.

[12] www.normemma.com/articles/armaslow.htm (accessed on Dec. 18, 2011)

[13] Nunes, P. and Kambil, A. (2001), Personalization? No Thanks., Harvard Business Review; Apr2001, Vol. 79 Issue 4, pp. 32-34

[14] Jupitermedia Corp. (2003), Beyond the Personalization Myth: Cost-effective Alternatives to Influence Intent

[15] Jarvis, J. (2009), What Would Google Do ?, HarperBusiness; 1st Edition

[16] Ho, S.Y., Bodoff, D. and Tam, K.Y. (2011), Timing of Adaptive Web Personalization and Its Effects on Online Consumer Behavior, Information Systems Research; Sep2011, Vol. 22 Issue 3, pp. 660-679

[17] Sull, D. (2009), How to thrive in turbulent market, Harvard Business Review; Feb2009, Vol. 87 Issue 2, pp. 78-88

[18] Liker K.J. (2004), The Toyota Way, McGraw-Hill; 1st Edition

[19] Turban, E., Lee, J.A., King, D. , Liang, T.B. and Turban, D. (2010), Electronic Commerce a Managerial Perspective, Pearson Prentice Hall; 6th Edition

[20] Miller, R. (2008), That is customer focus, BookSurge Publishing

[21] Hess, E. (2010), Smart Growth, Columbia Business School Publishing; 1st Edition

 ISBN: 978-1-61208-176-2

[22] Resinck, R. (2010), Fly Higher, Entrepreneur; Sep2010, Vol. 38

[23] Becker, M. and Arnold, J. (2010), Mobile Marketing For Dummies, Wiley Publishing, Inc. ; 1st Edition

[24] Dushinski, K. (2009), The Mobile Marketing Handbook, Information Today, Inc.; 1st Edition (January 19, 2009)

 ISBN: 978-1-61208-176-2

New Service Development Method for Prosumer Environments

Ramon Alcarria, Tomás Robles, Augusto Morales, Sergio González-Miranda
ETSI Telecomunicación
Technical University of Madrid
Madrid, Spain
{ralcarria,trobles,amorales,miranda}@dit.upm.es

***Abstract*—Prosumer environments are characterized by user participation in the service creation and provision processes. These services, which become increasingly important in recent years, have some peculiarities that differentiate them from the services that follow the traditional model of supplier-customer. However, there is little research on how to adapt existing business models to harness the prosumer's value and the implications of this new role for the company's business model. In this paper we design a methodology for the development of prosumer services by using the New Service Development approach to provide creation tools, used by prosumers to create final services. We pay special attention to the relationship between creation process participants by modeling this relationship as co-creation mechanisms. The proposed method is applied to a use case, based on prosumer interaction in the Future Intelligent Universe.**

Keywords-new service development; co-creation; service composition; service customization; QFD

I. INTRODUCTION

A new service provision model is needed in a society in which individuals, companies and cities are related; and in which users contribute with his suggestions, interesting information and even his own services to the rest of the community. This paper focuses on a new environment for the current society, based on user participation, not only in information provision but also in the creation and composition of their services, called *Prosumer Environment*.

Users participating in this society, called *prosumer users* or *prosumers*, want to get involved in service development stages, but they are not experts in the use of traditional tools of service development. Companies are aware of the evolution of the society and they view their customers as important resources when they develop new products and services. Thus, they try to involve their customers in the co-creation and co-development of new services [14]. Currently, some projects focused on the figure of prosumer are appearing, such as iStockphoto or Lego Mindstorms, and the term prosumer begins to be used by companies such as Sony to describe video camera users, producers and publishers of multimedia content.

From our knowledge, there are no product and service development models that explicitly take into account the new prosumer role but only as customer involvement in the business process [15] [4]. Nor is there much information on the related work on how to adapt existing business models to harness the prosumer business value and the implications of this new role for the company's business model.

The New Service Development (NSD) methodology is often used for the development of services that are new to the company, and that involve resources, processes and customer interaction [2]. In this paper, we extend the development model proposed by NSD in order to involve prosumer users, who wish to take responsibility for the creation and provision of services, in the service development process. The benefits of this new business model is perceived from the viewpoint of the company, which harnesses the power of the prosumer development to improve and test new products, and from the prosumers' point of view, which get involved in service development by using creation tools adapted to their experience level.

The rest of the paper is as follows. Section II describes the service provision prosumer model we want to achieve with our method and Section III analyzes and discusses the related work regarding service development with user participation. Section IV describes our proposed method through the NSD stages. Section V presents a scenario in which this methodology has been applied as a validation and Section VI presents the conclusions of our work.

II. PROSUMER MODEL

Internet has become a powerful distributed infrastructure that enables information to be widely available and its actors to interact with the rest of the world. Users require tools for providing their own services and consuming services published by others, and thus, transform the network into a collaborative infrastructure, adapted to different areas, such as social, personal and commercial ones. There are some initiatives to provide this type of tools, such as ICT-2007.1.6 challenges (14 projects) of the European FP7 program, focused on validating highly innovative and revolutionary ideas for new service paradigms.

The term *prosumer* [16] (as an acronym formed by the fusion of the words producer and consumer) is applied to those users that are at the same time consumers and producers of services or contents. The proposed prosumer model is shown in Fig. 1 and is described below. In the creation process, the prosumer, using his mobile phone or a computer, can design his services using the tools he has at his disposal. This is the most critical stage, because of the technical difficulty of transforming service creation ideas from a non-expert user to machine interpretable code. The *Creation tools* are oriented to a specific personal or

 ISBN: 978-1-61208-176-2

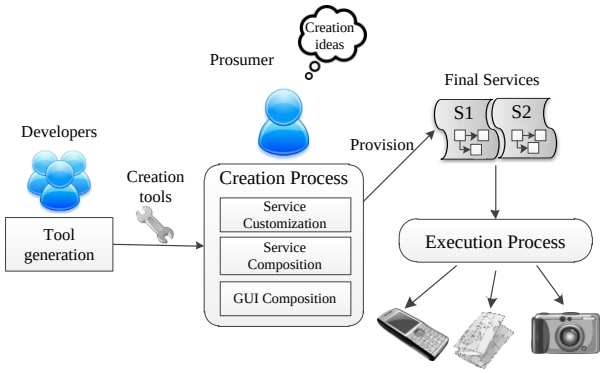


Figure 1. Prosumer model

professional domain. They are provided by companies or third-party developers, using the traditional software development paradigms. *Service composition* graphical environments have been developed to solve this problem, using different paradigms (automatic, semiautomatic, static workflow-based service composition [20], natural written language [19], etc.). *Customizable services* are service models that already solve most of the technical issues regarding service creation and they allow the prosumer to introduce some configuration and customization parameters, both in GUI and in the service logic. From the service presentation's point of view, the service composition paradigms that have proven to be more effective are based on User Interface Composition as some Mashup contributions [18] or those based on predefined templates [17]. The former allow greater flexibility in creation process while the latter usually have a more eye-catching look, providing that they have not been assembled by individual graphical blocks. The result of this stage is called *final service*, completely or partially created by a prosumer user and ready to be consumed.

Service provision and execution take place after the creation process. As these stages are very similar to those of traditional services we will not explain them in detail.

The complexity in the creation phase requires an infrastructure deployment to facilitate contact between the developers of creation tools and the prosumer user that will use these creation tools. A methodology is required to decide between different design strategies, according to the captured design requirements. Section IV explains the method we developed.

III. Related Work on Service Development

In this section, we focus on the related work about two key aspects developed in this paper: User participation in service development process and, specifically, the relationship between users and other participants in the NSD methodology.

A. Customer participation in service development

In recent years, companies have considered the importance of the presence of the customer in service development in order to understand customers' needs and wishes properly [21], evolving from the traditional model of service development, which produces common failure to involve service personnel and customers [1]. The presence of the customer in the production process results in increased customer value, as the overall benefit perceived in the solution at the price the customer is willing to pay [22]. Customer value is an aspect of the service that must be continually revisited for the company so that it can anticipate an alteration in customer's needs (the customer's perspective on a service offering can change from being favorable to being unfavorable).

With the emergence of the Web 2.0 in the information and communications society, based on user participation, the user acquires a leading role in service development. Therefore, user involvement in the service production process is more justified. Some authors believe that users are the root of the service idea. Matthing et al. [23] illustrate that the consumers' service ideas are found to be more innovative, in terms of originality and user value, than those of professional service developers.

The role that the user plays is evolving from "content prosumer", who uses the Internet and other technologies to find information and also to produce content, to "service prosumer", which develop services and make them available to other users. The FIA (Future Internet Assembly) mentions this evolution of Internet users in his roadmap [27], defining prosumer as "a new kind of Internet user, playing both roles consumers of services as well as creators of added value services based on those consumed". In this paper we focus on this second type of prosumers.

B. Co-creation in New Service Development

NSD is a service development methodology that is often used in corporate environments [2][21][22]. Johnson et al. synthesized past service development research and created a general four-stage NSD process model involving the phases of design, analysis, development and full launch, emphasizing the interdependence on design and development as well as the cyclical aspects of the new service creation process [2]. The difficulty of finding flexible tools for creating prosumer services is covered with large number of tools that relate the prosumer with NSD. In this section we review the tools, methods and practices found in the literature.

In the design phase, related tools focus on identifying costumer needs. Although there are tools that extract qualitative information on customer perception, as focus groups and face-to-face interviews, and stimulate the production of new ideas (brainstorming), other authors [6] consider that the best method to identify customer needs is to select the so called "lead users", which present strong needs. In our work, we define *domain-expert* as a lead user with good knowledge of his environment and who is aware of specific and general needs of the users it represents. Data mining techniques (artificial neural networks, decision trees, case-based reasoning, and multivariate discriminant analysis) [7] are used for classifying user need types for recommendation systems.

In the analysis phase, recommendation systems use techniques to analyze the information on user needs,

 ISBN: 978-1-61208-176-2

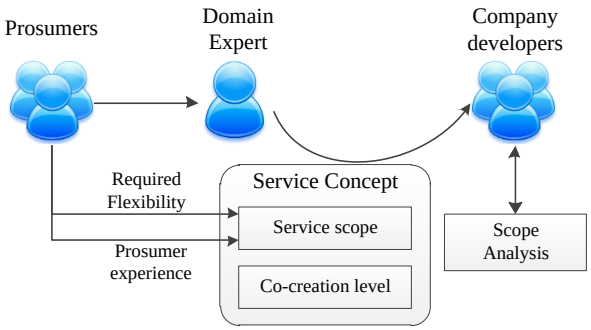


Figure 2. Design phase

processed in the previous stage in order to quantify its importance and thus, estimate the value of service for the target customers and the company. One of these techniques, the conjoint analysis [8], consists of a multivariate statistical technique which studies the joint effects of service components on consumers. This technique is often used along with the graphic technique of perceptual mapping [8], which helps marketers to visually display the perceptions of customers or potential customers.

Related to service development, Quality Function Deployment (QFD) [4] tools have been proposed to transform the needs of users into service design requirements, which are more easily understandable by developers. Once the service concept has been created, other service engineering models are used, which map, describe, and/or analyze the design of service processes and include the customer experience in form of interactions through the process [9]: blueprinting, SADT (Structured Analysis and Design Technique), STA (service transaction analysis) and IDEF3. Among these methods we highlight SADT, which proposes the involvement of NSD providers, project managers and customers.

Regarding the launching phase, the quality of the service being deployed and the customer satisfaction with the service once deployed is evaluated. Before service launching it may be necessary to identify potential failures in the service design or implementation. One of the most used techniques is the Failure Modes and Effects Analysis (FMEA) [10], which identifies failure modes based on past experience with similar products or services and provide corrective actions. The model proposed by Kano et al. [11] is often used, which complements the QFD to measure customer satisfaction and ranks customer demands with threshold attributes, in such a way that if a new service is not examined using the threshold aspects, it may not be possible to enter the market.

IV. PROSUMER NSD METHODOLOGY

In this work, we rely on the NSD model to define a New Service Development method in which the prosumer is present from service conception to the deployment of the infrastructure and the tools to personalize and provide services to other prosumers that consume them. The ultimate goal of this methodology is that users unfamiliar with traditional creation tools may be responsible for the creation of final services (through mechanisms such as composition or template customization).

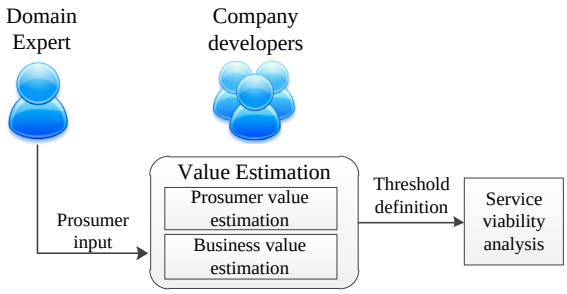


Figure 3. Analysis phase

Our methodology is based on the phase model proposed by Johnson et al. [2]. In this section, we describe our adaptation of the processes carried out at each stage in order to involve customers as service co-producers.

A. Design

As shown in Fig. 2, in this first stage we identify the target customer and choose a customer representative. We call him *domain expert* and he serves as a link between prosumers and company developers. To meet the prosumer involvement needs we have adopted a customer-centric approach and, using the technique of Quality Function Deployment (QFD) [4], we ensure that the dimensions most valued by customers are adequately captured and translated into objective metrics, as the main task of service development is to create the right prerequisites for the service [1]. To do this, the domain expert must acquire the best possible knowledge of the target customers (prosumer users) to feed QFD, and he should be able to appropriately translate customer expectations into design requirements.

A major challenge in this stage is to ensure that every decision is made based on delivering the correct services to potential customers. So, we first determine what type of service is going to be developed and the characteristics of the target customers. We define the *service concept* for prosumer environments as the combination of the *co-creation level* and *service scope*. The service concept helps to focus the relationship between customer needs and the company's strategic intent. The co-creation level measures the domain expert involvement in the NSD process. A low co-creation level means that the users are little involved in the co-creation process. We define service scope as the set of service requirements, covering both the design (functional and non-functional requirements) and the co-creation. The service scope is affected by two attributes: level of *flexibility* demanded in the service and *experience* level of the target prosumers regarding service creation. These attributes are interdependent, so that a lack of experience using creation tools will involve the development of creation tools with a lower degree of flexibility, which inevitably affects the service scope. Likewise, a service which requires a wide scope demands a high degree of flexibility and experienced target prosumers. As the service scope is defined by the target users, or the domain expert that represents them, the design phase analyses whether the relationship between service scope, flexibility and experience level of target customers is met. If the analysis fails some solutions are

 ISBN: 978-1-61208-176-2

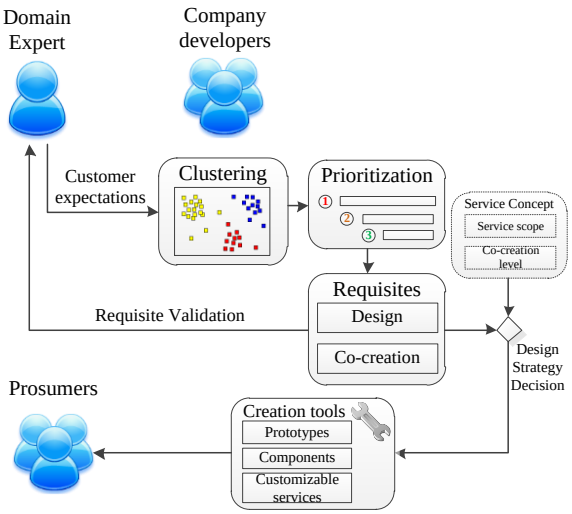


Figure 4. Development phase

sought, such as increasing user experience through training courses for prosumers or reducing the service scope by dividing it into several sub-services, whose combined scope makes up the full scope.

B. *Analysis*

In the analysis stage, organizations estimate the market-performance potential, strategy and financial prospects of the new service concept. The analysis phase, shown in Fig. 3, analyses whether the new service concept is aligned with the business strategy. Two simultaneous value calculation processes are produced, which determine the suitability of the service creation process: prosumer value and business value estimations.

We define *prosumer value* as the overall benefit perceived in the service solution at the price the prosumer is willing to pay. In the prosumer environment, users create and consume services as long as their perception (i.e., the prosumer value) of the service creation and provision tools is adequate. Prosumer value must be identified at early stages of the methodology by using customer input information. While customers may be able to provide some guidance, some of the most successful cases of value identification occur when a company provides a service that addresses a need that the customer was not aware of previously [3].

We define *business value* as the benefit experienced by the company for the acquisition of knowledge, experience and presence in the sector. The business model of prosumer service development has been exploited before, and is adopted by major online stores such as Android Market or Apple Store. These stores provide development tools, a service search and publication infrastructure and a control mechanism for published applications. In return, they benefit from a percentage of the applications' selling price, which is estimated by the application creator.

We define a threshold below which the project is not viable. This threshold depends on the benefit margins of the project development and the business value mentioned above. These concepts will not be studied in depth in this paper.

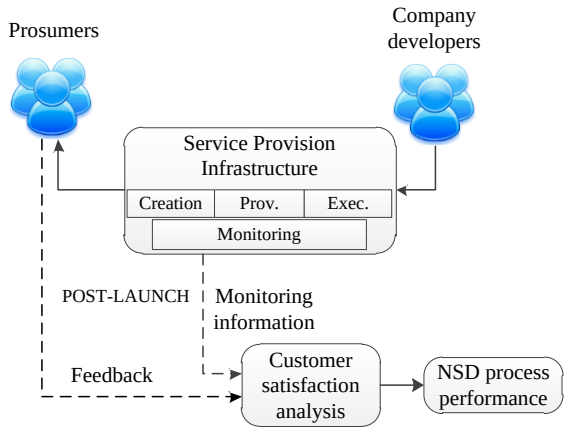


Figure 5. Launch phase

C. *Development*

The development phase is described in Fig. 4. In the model proposed by QFD, customer expectations are often described as verbations, for example: "I want to have an on/off button in the main screen of my creation tool". At this stage these verbations must be converted to re-worded data using simple expressions as "higher customization" and "add control buttons". These data are *classified into clusters* that share a common theme (e.g., interface customization, easy to use) and *prioritized*. The *house of quality* diagram [13] can be used for this task, which utilizes a planning matrix to define the relationship between customer desires and firm capabilities. The result of this classification and prioritization is a set of requisites (design and co-creation requisites). If the domain expert involvement is high it is advisable to validate with the domain expert the collection of customer expectations to meet the term *voice of the customer* [12], widely used in QFD to describe the in-depth process of capturing customers preferences and expectations. Once the requirements are validated, decision criteria are used to select *service design strategies*. The attributes affecting the service concept are used as criteria: *service flexibility* of prosumer tools demanded by customers and *experience level* in service creation for target customers. Fig. 6 shows the relationship between these two attributes and characteristics of the design strategies for the described use case.

A process whose objective is to develop a co-creation support system takes place once the design strategies are chosen. This system allows customers and developers to agree upon prototype evaluation, intermediate versions and co-creation requirements adjustment. Prototypes are used as proof of concept to demonstrate that the requirements of the prosumer are properly understood by the creation tools developers. Prototyping is considered a design strategy for environments with high involvement of the domain expert.

The outcome of this phase is a set of *creation tools*, templates, prototypes, atomic and customizable services as well as the mechanisms for monitoring the service life cycle.

D. *Launch*

The launch phase (see Fig. 5) is divided into two phases.

 ISBN: 978-1-61208-176-2

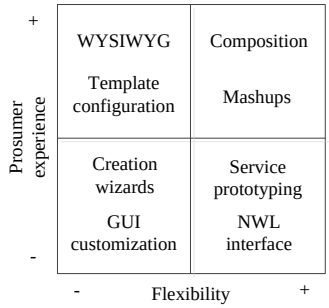


Figure 6. Design strategy decision criteria.

In the *pre-launch analysis* phase, the service provision infrastructure is designed, according to the developed services and the mechanisms for monitoring them.

In the *post-launch evaluation* phase, the information collected by the infrastructure and customer feedback is iteratively analyzed in order to evaluate the *prosumer satisfaction* for the service (NSD outcome quality). It also evaluates the process performance for the company (NSD process performance), from the point of view of the operational effectiveness and market-place competitiveness [5]. Measures to analyze this performance are divided into the prosumer satisfaction by the use of developed services (*NSD outcome quality*) and the process performance for the company (*NSD process performance*).

V. VALIDATION CASE

We present a scenario in which we applied this methodology for the creation of prosumer service delivery infrastructure in the Future Intelligent Universe, as part of the mIO! project [26], supported by the CENIT Spanish National Research Program (CENIT-2008 1019).

In this scenario the prosumer user, using his mobile phone, interacts with the elements (sensors, actuators, smart devices) that are around him in order to obtain their functionality. The prosumer user creates and shares services by following the prosumer model shown in Fig. 1. Our goal is to use our NSD method, defined in Section IV, to create an infrastructure for service creation, delivery and execution.

The first step is to select a domain expert (based on the concept of lead user [6]) to identify customer needs. We determine that the user experience is varied, because, although this environment is focused on non-expert users, some of them are familiarized with creation paradigms. Thus, if we want to achieve high flexibility, the scope analysis requires dividing the infrastructure design in several subsystems, to cover the service scope.

The viability analysis concluded with satisfactory results, thanks to high prosumer value and a sustainable business model for both the company and prosumers. The company business model is based on sponsored final services and the incorporation of advertisings and sponsorship in creation and execution tools. The business model for prosumers is based on applying the *Freemium model*, combination of free and premium services.

In the development stage, following the proposed QFD model, the ideas expressed by the domain expert are processed and clustered. In total we extracted 104 requirements, divided into 9 clusters (Infrastructure, User and Subsystem Interaction, Security, Context, Personalization, Service Model, User Interface and Technologies) y 32 subclusters. For example, through this process, the user idea "I want to take my services on my mobile" was turned into the "mobile service provision" requirement, and was introduced in the subcluster called *Service Provision*, within the *Service Model* cluster.

We relate the prosumer experience level to the creation flexibility required in order to determine the most appropriate design strategy. Fig. 6 shows the considered decision strategies.

As prosumers in this scenario have a mixed experience for service creation we need to take more than one design strategy. The main determinant for the decision is the requirement of high flexibility. Thus, we discard the strategies that provide only a few customization options (*GUI customization*, *Template configuration*, *Creation wizards*), and also the *WYSIWYG* (What You See Is What You Get) paradigm, for being difficult to be used in a domain as general as mobile and ubiquitous sensor access. Regarding the remaining options, we discard the *Mashup* creation strategy for being somewhat less flexible than the choice of *Service composition*. Therefore we consider the development of tools to enable *Service composition* for experienced prosumers and Natural Written Language (NWL) creation interface for non-expert users. Thanks to a high domain expert involvement in NSD we consider advisable to use *Prototyping* as a strategy to avoid deviations in the design objectives.

After selecting the design strategies we proceed to the infrastructure implementation stage. Some parts of the implementation, such as service creation environment based on natural written language and sensor access can be seen in [24] and [25] respectively.

The launch phase is performed on a test group that studied the platform and we received several conclusions. Due to space limitation we only describe four of them:

- The creation mechanism for service composition is somewhat difficult to assimilate by non-experts, who have found the creation system based on natural written language [24] more intuitive.
- To really obtain consumer satisfaction we need to provide a large set of components so that composition is versatile.
- The limited display capabilities of a mobile terminal delegate the composition creation method to devices with larger screens, such as tablets.
- It is recommended to support service provision with the help of fixed infrastructure, to avoid provision issues due to lack of coverage or battery problems.

As conclusions of the analysis we believe that the application of our NSD method to the implementation of this scenario has been very helpful and the application of the defined design strategies has been successful.

 ISBN: 978-1-61208-176-2

VI. CONCLUSION AND FUTURE WORK

This paper proposed a tool development method for creating services for prosumer users, based on NSD. We evolve this methodology to cover the interrelationships of the different roles that traditionally participate in the creation process (company developers and customers) with the new *prosumer user* role. This prosumer user is the consequence of the evolution of information society, in which users are more participative and feel responsible for the generation of services and their publication into the user community.

In this paper, we develop a NSD method for prosumer environments. We review some concepts belonging to the service development process, such as *service concept*, *service scope*, *value estimation*, and we define other concepts, as *co-creation level*, *prosumer value* and *co-requisites*, which became contributions to the traditional NSD stages. As a validation, we show a service creation scenario for mobile prosumers and, following our proposed method, we develop some creation tools and a service delivery infrastructure.

We consider the implications in NSD of the new prosumer role as a contribution to the traditional service development process. This prosumer role reflects the current evolution of society towards a service provision model focused increasingly on the user.

Feedback from users allows us to identify future work on the proposed model, such as developing service composition tools for distributed service provision and task delegation and studying the characteristics of creation skills of non-expert prosumers thoroughly.

REFERENCES

[1] A. Johne and C. Storey, "New Service Development: A Review of the Literature and Annotated Bibliography," European Journal of Marketing, 1988, vol. 34, pp. 184-251.

[2] S.P. Johnson, L.J. Menor, A.V. Roth, and R.B. Chase, A critical evaluation of the new service development process: integrating service innovation and service design. Eds. Sage Publications, Thousand Oaks, CA, 1999, pp. 1-32.

[3] C. Liu, "Constructing a value-based service development model," Journal of Applied Business Research, Dec. 2006, vol. 22, pp. 47-60.

[4] T. Ohfuji and T. Noda, Quality function deployment: integrating customer requirements into product design. Eds. New York: Productivity Press, 2004.

[5] M.V. Tatikonda and M.M. Montoya-Weiss, "Integrating operations and marketing perspectives of product innovation: the influence of organizational process factors and capabilities on development performance," Journal of Management Science, 2001, vol. 47, no. 1, pp. 151–172.

[6] E. von Hippel, "Lead users: A source of novel product concepts," Journal of Management Science, 1986, vol. 32, pp. 791-805.

[7] K. Kim, "Customer Need Type Classification Model using Data Mining Techniques for Recommender Systems," World Academy of Science, Engineering and Technology, vol. 80, pp. 279- 284

[8] K. Ramdas, O. Zhylyevskyy, and W.L. Moore, "A Methodology to Support Product-Differentiation Decisions," IEEE Transactions on Engineering Management, Nov. 2010, vol. 57, no. 4, pp. 649-660.

[9] J. Frauendorf, Customer Processes in Business-to-Business Service Transactions. Eds. duv, Nov. 2006.

[10] Z. Bluvband and P. Grabov, "Failure analysis of FMEA," Reliability and Maintainability Symposium, Jan. 2009, pp. 344-347.

[11] N. Kano, N. Seraku, F. Takahashi, and S. Tsuji, "Attractive quality and must-be quality," The Journal of Japanese Society for Quality Control, 1984, vol. 41, no. 2, pp. 39–48.

[12] E. Roman, Voice-of-the-Customer Marketing: A Revolutionary 5-Step Process to Create Customers Who Care, Spend, and Stay. Eds. McGraw-Hill, Sep. 2010.

[13] D. Kim, "Application of the HoQ framework to improving QoE of broadband internet services," IEEE Network, March-April 2010, vol. 24, no. 2, pp. 20-26.

[14] L. Witell, P. Kristensson, A. Gustafsson, and M. Löfgren, "Idea Generation: Customer Cocreation versus Traditional Market Research Techniques," Journal of Service Management, 2011, vol. 22, no. 2.

[15] A. Lundkvist and A. Yakhlef, "Customer Involvement in New Service Development: a conversational approach," Managing Service Quality, 2004, vol. 14 no. 2/3, pp. 249-257.

[16] A. Toffler, The Third Wave. Eds. Bantam Books, 1980 .

[17] J. Ling, P. Ping, Y. Chun, L. Jinhua, and T. Qiming, "Rapid Service Creation Environment for service delivery platform based on service templates," in Proc. of IFIP/IEEE International Symposium on Integrated Network Management, June 2009, pp. 117-120.

[18] Q. Zhao, G. Huang, J. Huang, X. Liu, and H. Mei, "A Web-Based Mashup Environment for On-the-Fly Service Composition," in Proc. of IEEE International Symposium on Service-Oriented System Engineering, Dec. 2008, pp. 32-37.

[19] M. Cremene, J-Y. Tigli, S. Lavirotte, F-C. Pop, M. Riveill, and G. Rey, "Service Composition Based on Natural Language Requests," in Proc. of IEEE International Conference on Services Computing, Sept. 2009, pp. 486-489.

[20] N. Laga, E. Bertin, and N. Crespi, "User-centric Services and Service Composition, a Survey," in Proc of Annual IEEE Software Engineering Workshop, Oct. 2008, pp. 3-9, 15-16.

[21] B. Edvardsson and J. Olsson, "Key concepts for new service development," The Service Industries Journal, April 1996, vol. 16, no. 2, pp. 140-164.

[22] M. Reinoso, S. Lersviriyajitt, N. Khan, W. Choonthian, and P. Laosiripornwattana, "New service development: Linking resources, processes, and the customer," in Proc, of Portland International Conference on Management of Engineering & Technology, Aug. 2009, pp. 2921-2932.

[23] J. Matthing, B. Sanden, and Bo Edvardsson, "New service development: learning from and with customers," International Journal of Service Industry Management, 2004, vol. 15, pp. 479-498.

[24] G. Sebastian, J.A. Gallud, and R. Tesoriero, "A Proposal for an Interface for Service Creation in Mobile Devices Based on Natural Written Language," in Proc. of the Fifth International Multi-conference on Computing in the Global Information Technology, Sept. 2010, pp. 232-237.

[25] U. Aguilera, A. Almeida, P. Orduña, D. López-de-Ipiña, and R. de las Heras, "Continuous service execution in mobile prosumer environments," in Proc. of Int. Symposium of Ubiquitous Computing and Ambiente Intelligence, Sept. 2010, pp. 229-238.

[26] mIO! project Web page: http://www.cenitmio.es (web in Spanish)

[27] Future Internet Assembly Research Roadmap v1.0. Available at the European Future Internet Portal: http://www.future-internet.eu [Retrieved: Nov. 22, 2011]

 ISBN: 978-1-61208-176-2

Digital Investigations for Enterprise Information Architectures

Syed Naqvi, Gautier Dallons, Christophe Ponsard

Centre d'Excellence en Technologies de l'Information et de la Communication (CETIC)
29 Rue des Frères Wright, 6041 Charleroi, Belgium
{syed.naqvi; gautier.dallons; christophe.ponsard}@cetic.be

***Abstract*—This paper highlights the role of digital forensics in the enterprise information architecture. It presents a framework for embedding digital forensics analysis techniques at various stages of corporate information and communication technologies (ICT) lifecycle. A set of best practices for the corporate ICT security policy is also outlined to keep the operational costs of digital forensics at the optimal level. It also presents a detailed analysis of the risks to the competitive-edge of the companies that will not employ the forensics solutions to protect their business interests. This work also provides a high-level roadmap for the adaption of digital forensics in the emerging core business technologies such as cloud computing and virtualization infrastructures.**

Keywords - digital forensics analysis, ICT security architecture, enterprise information architecture

I. INTRODUCTION

Crimes involving computers started to occur as soon as the computers started sprawling across the various activities of everyday life in the 1980s. United States Federal Bureau of Investigation (FBI) launched its *Magnetic Media Program* [1] to address the growing needs of analyzing computers especially storage media for their investigations of high-technology crimes. The growth of networking technologies in the 1990s gave further impetus to the need for tackling crimes using sophisticated technological means. However, the overall scope of the digital forensics remained confined to the law enforcement agencies with the sole objective of collecting reliable evidences for the prosecution of the criminals involved in a court of justice.

The interest of performing digital forensics analysis at the enterprise level has significantly grown nowadays. This trend can be seen as a logical evolution of *white hat* or *ethical hacking* that is performed by the businesses to ensure that their ICT infrastructure is free from vulnerabilities. However, the major driving force behind this trend is the fact that growing number of *criminal offenses* using ICT is overwhelming the computer crime units. It is not so easy to involve law enforcement agencies in a commercial environment where there are some suspicious activities; but no explicit computer crime is committed. Moreover, their involvement casts shadows on the business interests notably on the reputation of the company.

Nowadays, enterprises are embracing considerable shift in their business approach where they have to not only adapt to non-traditional concepts such as virtualization, but also switch to more comprehensive security solutions to cope with the new security requirements as only pre-incidence measures, i.e., attack preventions is no longer be sufficient to protect their assets. They need to invest in the post-accident situation, i.e., digital forensics. It is evident that European enterprises can learn a lot in this area from their US counterparts who have acquired considerable experience of using digital forensic technologies in the business environment whereas digital forensics is still seen by a vast majority of Europeans as specialized tools for police to tackle cyber criminology.

This nascent trend of in-house digital forensics analysis in a relatively non-criminal context has to prove its worth by providing some competitive edge to the businesses. Major challenges include the demarcation of exact role of the digital forensics in the overall corporate ICT security operations; reducing its functional costs both in terms of monetary expenditures incurred in the acquisition of corresponding technologies; and in terms of the time consumed in the subsequent investigations especially when the operations of core business line are directly affected. This paper pragmatically identifies the role of digital forensics in the overall security architecture of an enterprise. We propose to implant the corresponding digital forensics analysis features in the various components of the corporate ICT infrastructure. This scheme provides better organization of the security functionalities with minimal impact on the overall performance of the ICT operations.

This paper presents a framework of digital forensics for enterprise information architectures and evaluates the scope of various methodologies and tools for these enterprise applications. It also presents a detailed analysis of the risks to the competitive-edge of the companies that will not employ the forensics solutions to protect their business interests. This work also provides a high-level roadmap for the adaption of digital forensics in the emerging core business technologies such as cloud computing and virtualization infrastructures.

This paper is organized as follows: Section II describes the role of digital forensics in the enterprise information architecture. A digital forensics analysis framework for enterprise information security policy is presented in Section III. Section IV highlights the impact of digital forensics on the competitive-edge of enterprises. Section V identifies a range of challenges of investigating the emerging paradigm of virtualized infrastructures. A discussion on the recent trends in the area of digital forensics and their positioning with our approach is presented in Section VI. Finally, some conclusions are drawn in Section VII.

 ISBN: 978-1-61208-176-2

II. ROLE OF DIGITAL FORENSICS IN ENTREPRISE INFORMATION ARCHETCTURE

The area of digital forensics analysis is considerably mature in the modern criminology; however, its scope in the commercial environment is still vague. Therefore, it is needed to precisely identify the role of digital forensics at various levels of the enterprise ICT operations so as to enhance the overall quality of protection without causing any substantial impact on the system's performance. This section outlines major blocks of a typical enterprise ICT infrastructure followed by the identification of those blocks where digital forensics analysis has a role to play. These roles are elaborated for individual blocks. These roles can eventually be harnessed together to constitute various digital forensics analysis functions.

Figure 1 shows a typical lifecycle of corporate ICT operations. These blocks are tagged to facilitate their subsequent usage. These operational blocks are subject to a number of internal and external requirements that have to meet in order to ensure the proper functioning of the overall business routines. We identify those blocks that require digital forensic analysis features for the improvement of the quality of protection offered by the corporate ICT security architecture. A magnified glimpse of these blocks is shown in the Figure 2.

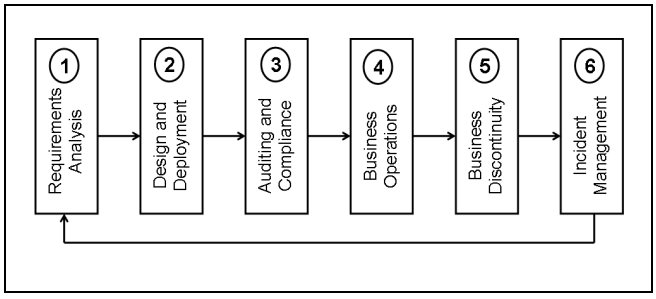


Figure 1. Various phases of corporate ICT operations

1) Design and Deployment

The design and development phase should ensure that the ICT infrastructure is globally *forensic-friendly* – this implies that both design specifications and deployment parameters should provide some sort of explicit checkpoints to facilitate digital forensics analysis at any time instant without causing major disruption to their operational routines.

The design and development phase also needs to ensure that *chronological documentation (chain of custody)* can be conveniently produced if deemed necessary. It can include some nonrepudiation techniques such as watermarking to justify that they are not tampered at any stage.

The design and development phase should also include *resilience planning* so that the forensic analysis will have minimal impact on the functioning of routine operations. This feature is particularly important for today's large-scale highly connected systems, such as Clouds, where disruption of entire ICT infrastructure in the aftermath of any incident will inflict huge damages. This feature also includes recovery time that is ideally kept at the minimum possible level.

2) Auditing and Compliance

The auditing and compliance is not only a prerequisite for launching a specific business but also a marketing tool to increase the customer base. Therefore, the auditing and compliance phase should provide means of justifying that the business is meeting all of its *legal obligations*. For example, the United States Sarbanes Oxley Act (SOX) [2] requires a formal process of using forensic analysis techniques for the investigation of incidents. This law has made significant impact on the security policies and incident management strategies of US based corporations [3].

Besides serving the compliance issues of the legal requirements, digital forensics can also play its role in the fulfillment of regulatory requirements that most of the corporations are required to comply with. For example, information security incidence management procedures are recommended by the ISO standard ISO27002 [4].

Likewise, the *quality assurance* practices can be improved by using digital forensics techniques such as analyzing the performance bottlenecks of network bandwidth, storage mediums, etc. They can be employed when some performance degradation is reported or they can be proactively used to ensure the execution of an optimal quality assurance plan.

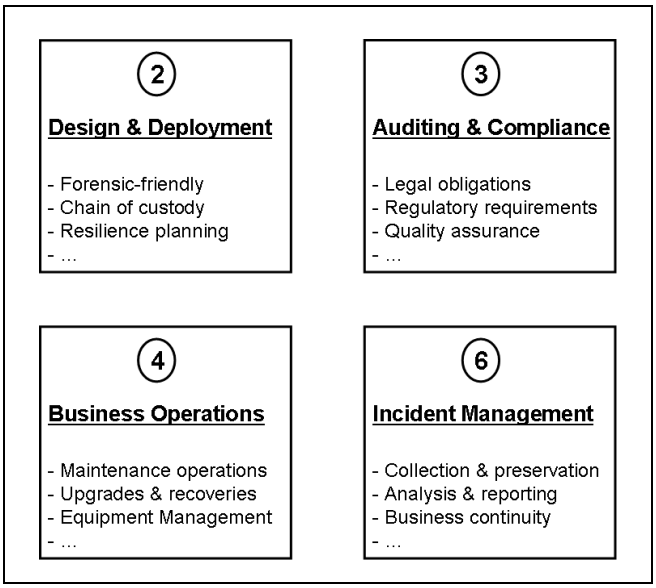


Figure 2. Corporate ICT phases that can be improved by Digital Forensics

3) Business Operations

The business operations phase is crucial for meeting the company's targets within the allocated resources. Therefore, it is important to explore newer techniques to improve this phase not only to facilitate the targets achievements, but also to provide a competitive-edge in this core operational phase.

Several *maintenance operations* of corporate ICT infrastructure, such as disk cleaning operations, require reliable analysis solutions. Digital forensics analysis tools

 ISBN: 978-1-61208-176-2

can be employed to efficiently provide these maintenance operations.

With the ever increasing evolution and expansion of the ICT scope in everyday life, the *upgrades* have become a routine activity. Moreover, *data recovery* from the obsolete, faulty, or broken equipment requires some reliable solutions. Digital forensics techniques can be employed for recovering data in these situations.

Managing the ICT inventories in a corporate environment also requires efficient sorting solutions before discarding the useless lot. Digital forensics techniques can assist in securely disposing of equipment.

4) Incident Management

This is the privileged phase for employing digital forensic techniques as the general perception of the *forensics* is the post incident analysis. Digital forensics can be used by the administrator of a corporate ICT infrastructure to gather the digital traces and consequently analyze them to determine the causes of incident. All these actions should be carried out within the given legal framework. Another important role of the incident management team is to ensure business continuity ideally even during the incident phase. In real terms, it should at least ensure the continuation of core business operations during the adverse situations.

III. A DIGITAL FORENSICS ANALYSIS FRAMEWORK FOR ENTERPRISE INFORMATION SECURITY POLICY

Corporate security policy can play the pivotal role in outlining best practices for their ICT security teams. We suggest that organizational security policy of a modern enterprise should explicitly reflect the role of digital forensics as described in the section II of this paper.

Figure 3 presents a placement of *security and forensics team* in a corporate ICT infrastructure where it interacts with the different phases that can benefit from various digital forensics analysis techniques.

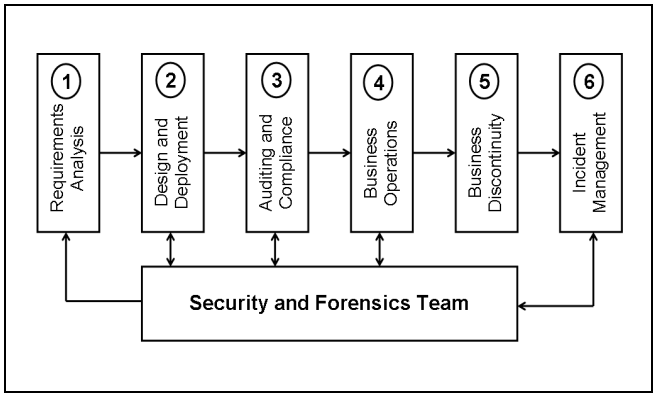


Figure 3. Security and Forensics Team in a Corporate ICT Infrastructure

It is essential that security and forensics team should be equipped with *fine-grained forensics policy* that can handle the peculiar requirements of the contemporary ICT infrastructures especially their complexities, scale, and decentralization. It is understood that highly dynamic environments such as *Open Clouds* give rise to enormous challenges of localization and demarcation of their boundaries that keep on changing in an amoeboid style. While research and development initiatives are needed to effectively address these specific challenges, ICT security teams of corporate sector can handle the digital forensics tasks by having clear vision and strategy of achieving the security goals. The current state of the tools support for the digital forensics activities is quite sparse. We therefore recommend using a set of tools for different tasks instead of looking for a comprehensive framework or toolkit that can provide a silver bullet kind of solution.

We now present different phases of standard forensics analysis and their efficient implementation in businesses. They are presented in the context of incident management; however, we have already shown that these techniques are equally useful in a number of other ICT operations.

A. Preparation phase

Preparation phase mainly involves adequate planning by envisioning the security threats and the contingency plans followed by the deployment of necessary tools and procedures. This is a management task that requires regular re-evaluation and assessment as the threats picture changes rapidly especially in the dynamic environments such as virtual infrastructures.

B. Detection phase

Detection phase often includes collection mechanism as well. This phase requires up to date set of tools to monitor the company's ICT resources and capture the various events followed by the identification of some abnormal activity that should be logged and the security manager be notified.

C. Preservation phase

Preservation phase generally involves the production of copies of the ICT resources being investigated. The original resources are preserved so that they can no longer be altered. The integrity of the original resources is indispensable in the follow-up procedures in the court of justice or even in the disciplinary action board.

D. Analysis phase

Analysis phase traditionally employs human interventions to note the sequence of events leading to a particular incident. However, the ever growing scale and scope of the digital data now necessitates the use of some semantic tools for the analysis of log files and traces. This phase also involves some artificial intelligent mechanisms for performing good quality analysis to avoid human errors.

E. Recovery phase

Recovery phase is generally not seen as an integral part of the forensics task. However, any emergency or incidence response cannot be completed if the system is not restored to its original functioning state. This situation is especially desirable when the core business of a company is halted due

 ISBN: 978-1-61208-176-2

to some incidence and therefore there is immense pressure to restore the routine activities.

F. Reporting phase

Reporting phase is not necessarily the preparation of legal case against the attackers. A simple report of incidence for the line manager can also constitute this phase. The objective is to produce a record of the incidence that took place in the corporate ICT infrastructure. It can also be used as a feedback for the preparation phase, so that the reasons leading to the incident can be taken into account by the company's security team.

IV. Impact of Digital Forensics on the Competetive-edge of Enterprises

There are genuine ICT security concerns for enterprises especially when the scope of virtualization brings several challenges for its deployment including a lot of uncertainty as to how and where to implement security [5]. Security and dependability issues are a gauging factor for measuring the success of business endeavors. Classical security solutions and practices are getting obsolete in the face of the peculiar security requirements of virtualization infrastructures where physical resources are dynamically mapped to address the spontaneous business needs. The inherent nature of virtualization requires totally different security provisioning approach than the classical one developed decades ago.

This section covers a set of threats mitigation techniques that can be employed by using digital forensics techniques and the risks of not using digital forensics in the enterprise environments. We maintain that it is becoming indispensable for the businesses to integrate digital forensics in the security architecture of their enterprise information architectures otherwise there could be devastating impact both in terms of security breaches and in the loss of business prospects.

A. Threats Mitigation Techniques using Digital Forensics

This section summarizes a set of major security challenges that can be addressed by using appropriate digital forensic technologies.

1) Access Control

Enterprise virtualization infrastructures offer promising features for the enterprises. However, virtual ICT resources are lot more vulnerable to malicious activities than the classical ones. Smart solutions for the monitoring of the access logs of an enterprise have become indispensable to know if its resources are the target of some malicious activity. It also includes insider threats such as frequent remote access to the company's resources outside the routine office hours.

2) Steganalysis

The ever-growing demand of bandwidth capacity for the contents rich applications is driving the conception of high speed internet backbone to enable seamless access to the applications using bulk of data. However, this capacity exposes enterprises to *steganography* where important corporate data can easily be stolen. The enterprises need to incorporate efficient steganalysis tools to protect its intellectual properties.

3) Multi-tenancy

Multi-tenancy is quite new concept [6] that refers to the architectural principle, where a single instance of the software runs on a software-as-a-service (SaaS) vendor's servers, serving multiple client organizations. Assuring data isolation on a node is a challenging task for the security designers that include complete isolation of their execution environments and data storage including temporary storage of the execution data. Enterprises need to employ some analysis tools to trace the software instances for being assured of the protection of their data in the multi-tenant applications.

B. Risks Analysis for Computing Impact on the competitive-edge of Enterprises

The ISO27001 [7] standard defines the way to establish an Information Security Management System. Considering this standard, a company has to monitor its security in order to adapt the security policy to new threats. This standard considers not only IT security but also Information system security that is broader than the IT infrastructure. In this paper, we only consider IT security of the infrastructure used to operate the enterprise information system.

ISO27001 requires enterprise to monitor its *security state* in order to monitor security breaches and to react accordingly. Security monitoring implies tracking of the security traces left by an attacker. Digital forensics analysis techniques and tools can facilitate this task. This standard requirement has the objective of reducing the risk linked to security for companies. If we consider a company with a set of computer business asset A = {a1, …, an}, each asset is concerned by security following four properties : Confidentiality (C), Integrity (I), Availability (D) and Legal force (L) (legal force is the ability of an asset to be used as a proof in the case of lawsuit). Each asset has some requirements concerning these properties. The risks are to loose one or more of these security properties. The risk of a particular enterprise can be summarized by $\sum$(Ic(ai)+ Ii(ai) + Id(ai) + Il(ai)), where Ic is the financial impact function of confidentiality lose, Ii is the financial impact function of integrity lose , Ii is the financial impact function of integrity, Id is the financial impact function of availability and Il is the financial impact function of legal force lose. We can argue that a company that doesn't use forensics techniques is exposed to a financial risk of $\sum$(Ic(ai)+ Ii(ai) + Id(ai) + Il(ai)). Now, the question is how digital forensics can reduce this risk. Digital forensics will be used as a curative tool. Forensics techniques will be used after the attack when an impact has been detected. In this scope, it will be only useful to demonstrate legally the cybercrime and to try to obtain compensation accorded by a court of law. If we consider our maximal financial risk will be reduced by an amount equal to the compensation obtained through a court of law; but some fees must be paid to the lawyers. We can

 ISBN: 978-1-61208-176-2

therefore express this risk by of ∑(Ic(ai)+ Ii(ai) + Id(ai) + Il(ai)) – compensations + fees. This financial risk formula is interesting:

- More ∑ Il(ai) is high, more compensations will be important due to difficulties to demonstrate the cybercrime
- If fees are higher than compensations, the forensics analysis won't be useful

So, optimizing the benefit of forensics investigation implies to reduce the risk of losing legal force of an asset and its traces (ideally ∑ Il(ai) must be closer to 0). It is also important to reduce the lawyers' fees by contracting a legal insurance. Legal insurance is marginal compared to cybercrime compensation. In this situation, our risk can be evaluated to ∑(Ic(ai)+ Ii(ai) + Id(ai)) – compensations.

Compensations are relative to security incident of some assets and are at less equivalent to direct damages. In this case, the security impacts of these assets are at least covered. The risk equation becomes $\sum_{i<>j}$ (Ic(ai)+ Ii(ai) + Id(ai)) - |delta| where delta is the difference between compensation and the impacts of assets aj.

This formula demonstrates that forensics reduces significantly the financial risk by suppressing the risk of the assets that can be fully covered by forensics. And an extra reduction is induced by compensations. In an ideal and idyllic situation, each asset can be traceable and all the legal demonstration of cybercrime can be produced in this case forensics can generate an extra benefit equal to |delta|.

V. CHALLENGES OF INVESTIGATING VIRTUALISED CORPORATE INFRASTRUCTURES

The concept of *virtualization* is not new in the field of ICT. It dated back to the inception of programming language compilers that virtualizes the object code [8]. However, the concept of *virtualization infrastructures*, where physical resources are dynamically mapped to address the spontaneous business needs, is relatively new. Moreover, the scale and scope of this novel concept brings several challenges for its deployment including a lot of uncertainty as to how and where to implement security [5]. The inherent nature of virtualization requires totally different security provisioning approach than the classical one developed decades ago.

Classical digital forensics techniques and solutions require precise information of the underlying infrastructure to perform investigation and conceive the sequence of events. They cannot be applied to these emerging infrastructures due to the intrinsic characteristic of virtualization that provides abstraction to the underlying resources and infrastructures. This section examines the challenge of investigating virtualized corporate infrastructures that will have to be addressed to ensure smooth and secure transition from the classical enterprise information architecture towards virtualized one.

A. Conducting Security Audit Investigation of Future Internet-based Virtualization infrastructures

Security audit assess the security of the networked system's physical configuration and environment, software, information handling processes, and user practices. While it is similar in terms of investigation, it is carried out before commissioning of a system and then on regular basis to ensure the desired functioning of a system. Whereas the investigations carried out by the digital forensics team is the post incident activity where a malicious activity successfully carried out its nasty action. However, in terms of *investigating* an ICT infrastructure, both have similar challenge of dealing with virtualization paradigm. We have therefore used our previous work on security audit [9] to leverage the work on applying digital forensics in the virtual infrastructures.

B. Case-study: Payment Card Industry Data Security Standard (PCI-DSS)

Various security audit standards such as Payment Card Industry Data Security Standard (PCI-DSS) require audit of the physical controls [10]. The virtualization infrastructures provide an abstraction layer to the underlying lower-level details. This situation raises several security concerns such as multitenancy; lack of security tools [11]; and disparity with the classical IT security audit practices.

There exist a number of generic monitoring tools such as hardware monitoring (e.g., HP Insight Manager, Dell Open Manage, VMWare Virtual Center, etc.), performance monitoring (e.g. VizionCore, Veeam Monitor, Vmtree, Nagios, etc.), machine state monitoring (e.g. Virtualshield, Logcheck, etc.), and security monitoring (e.g. intrusion detection, honeypots, etc.). However, these tools may not be suitable for security audit controls of virtualization infrastructures as physical controls can be distributed that will require onsite checks by the local controllers. There is a strong need of a new set of matrices for measuring security strength. With more reliable matrices, new check-pointing models need to be developed. Besides these technical requirements for carrying out security audit of the virtualization infrastructures, there is also a need of new regulations/legislations for the cross-border deployment of resources used in virtualization infrastructures. This work is the continuation of our previous work on analyzing the overall security requirements of deploying virtual infrastructures [12].

VI. DISCUSSIONS

The scope of digital forensics outside the criminology sphere was hardly explored in the past. However, the broadened scope of corporate ICT infrastructures and the reliance of core businesses on these infrastructures are pushing the paradigm shift in this domain. Some recent literature shows the exploration of digital forensics in the corporate sector [13,14]. However, these efforts are mainly focused on the philosophical possibilities. They do not

 ISBN: 978-1-61208-176-2

provide any concrete model or design approach towards the inclusion of digital forensics practices in the routine operations of corporate ICT infrastructures. Likewise, a set of best practices for computer forensics is proposed in [15] that describe some effective ways of carrying out the digital forensics analysis in the post-incident situations without any specific link to the commercial side of employing these techniques in the business environments.

There are the genuine predictions that near future cybercrimes will be driven by the *clouds* and virtualization infrastructures [16]. We believe that businesses and law enforcement agencies won't be able to cope with this wave of contemporary crimes with the classical digital analysis approaches. Digital forensics is often a very painstaking task that consumes enormous resources and takes considerable time to develop a sequence of events that is acceptable by the courts. An example is the FBI investigation of ENRON scandal [17] where FBI gathered and analyzed 31 terabytes of digital data from 130 computers; thousands of e-mails; and more than 10 million document pages. The entire investigation took five years while the total monetary costs incurred remained largely unknown. With the proliferation of computing through virtualization infrastructures and the evident increase of related crimes, the law enforcement agencies will simply be overwhelmed with the demand of providing digital forensic analysis requests. Therefore, there is a strong need for the businesses to include digital forensics in their corporate security strategy. They should use these technologies within the legal framework covering their activities.

We have carried out a risk analysis to show the impact of not using the digital forensic solutions by the enterprises. We proved that loss of clientele and sanctions from the regulatory bodies might severely harm the business interests of those enterprises who will not employ the digital forensics technologies for the protection of their business interests. A similar concern is reported in [18] that predict *fall of forensic research behind the market* in the next ten years if various disjoint research efforts are not harnessed together in a systematic way.

VII. CONCLUSIONS AND PERSPECTIVES

We have presented a non-classical approach towards the digital forensics analysis in this paper. We argue that enterprise information architectures can improve their overall quality of protection if digital forensics technologies are made their integral part. We presented a framework for embedding digital forensics analysis techniques at several stages of enterprise information lifecycle followed by a set of best practices for operating an enterprise information security policy together with the digital forensics techniques and tools.

There are a number of open issues that we plan to address in the near future. The foremost is the use of digital forensics solutions in the virtualization infrastructures such as *open clouds*. Major challenges of virtualization infrastructures are the absence of a fix perimeter of the ICT resources; and the unknown details of the underlying physical infrastructure.

ACKNOWLEDGMENT

The research leading to the results presented in this paper has received funding from the Walloon Region of Belgium through the project CE-IQS (Centre d'Expertise en Ingénierie et Qualité des Systèmes) and the European Union's seventh framework programme (FP7 2007-2013) Project PONTE under grant agreement number 247945.

REFERENCES

[1] K. S. Rosenblatt, High-Technology Crime: Investigating Cases Involving Computers, KSK Publications, ISBN 0-9648171-0-1, 1995

[2] The United States Sarbanes Oxley Act 2002 – http://uscode.house.gov/download/pls/15C98.txt <retrieved: Nov. 2011>

[3] J. Mullis, The Impact of the Sarbanes-Oxley Act of 2002 on Computer Forensic Procedures in Public Corporations, University of Oregon, July 2009 – https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/9480/Mullis-2009.pdf <retrieved: Nov. 2011>

[4] International Organization for Standardization, Standard ISO27002: Code of practice for information security – http://www.27000.org/iso-27002.htm <retrieved: Nov. 2011>

[5] R. Adhikari, The Virtualization Challenge, Part 5: Virtualization and Security, TechNewsWorld, March 2008

[6] F. Chong, G. Carraro, and R. Wolter, Multi-Tenant Data Architecture, Microsoft Corporation, June 2006

[7] International Organization for Standardization, Standard ISO/IEC 27001:2005 Information technology -- Security techniques -- Information security management systems -- Requirements

[8] J. Bloomberg, Building Security into a Service-Oriented Architecture, ZapThink Whitepaper, ZapThink LLC Publisher, May 2003

[9] S. Naqvi, G. Dallons, C. Ponsard, and P. Massonet, Ensuring Security of the Future Internet-based Virtualization Infrastructures (Position Paper), IEEE Symposium on Security and Privacy 2010, Oakland, CA, USA, May 16-20, 2010

[10] Payment Card Industry Data Security Standard (PCI-DSS) https://www.pcisecuritystandards.org/ <retrieved: Nov. 2011>

[11] E. Haletky, Virtualization Security – Security and Compliance within the Virtual Environment, DABCC online article 08 April 2009 http://www.dabcc.com/channel.aspx?id=279 <retrieved: Nov. 2011>

[12] S. Naqvi, P. Massonet, and J. Latanicki, Challenges of Deploying Scalable Virtual Infrastructures - A Security Perspective, CESNET Conference on Security, Middleware and Virtualisation, Prague, Czech Republic, September 25-26, 2008

[13] B. Nikkel, The Role of Digital Forensics within a Corporate Organization, IBSA Conference, Vienna, Austria, May 2006 – http://www.digitalforensics.ch/nikkel06a.pdf <retrieved: Nov. 2011>

[14] J. Heiser, Digital Forensics and Corporate Investigations, Gartner, November 2005 – www.gartner.com/teleconferences/attributes/attr_144863_115.pdf <retrieved: Nov. 2011>

[15] Scientific Working Group on Digital Evidence (SWGDE) Best Practices for Computer Forensics V1.0, 2004 – http://swgde.org/documents/swgde2005/SWGDE%20Best%20Practices%20_Rev%20Sept%202004_.pdf <retrieved: Nov. 2011>

[16] Trend Micro Report: The Future of threats and Threat Technologies – How the Landscape is Changing, December 2009 – http://affinitypartner.trendmicro.com/media/34716/trend_micro_2010_future_threat_report_final.pdf <retrieved: Nov. 2011>

[17] Federal Bureau of Investigations (FBI), Digital Forensics: It's a Bull Market, July 2007 – http://www.fbi.gov/page2/may07/rcfl050707.htm <retrieved: Nov. 2011>

[18] S. Garfinkel, Digital forensics research: The next 10 years, Elsevier Science Directi Magazine Digital Investigation vol. 7, pp 64-73, 2010 – http://www.dfrws.org/2010/proceedings/2010-308.pdf <retrieved: Nov. 2011>

 ISBN: 978-1-61208-176-2

Shadow IT

Management and Control of unofficial IT

Christopher Rentrop; Stephan Zimmermann
Faculty of Computer Science
HTWG Konstanz – University of Applied Sciences
Brauneggerstr. 55, 78462 Konstanz, Germany
rentrop@htwg-konstanz.de; stzimmer@htwg-konstanz.de

Abstract—**Shadow IT describes the supplement of "official" IT by several, autonomous developed IT systems, processes and organizational units, which are located in the business departments. These systems are generally not known, accepted and supported by the official IT department. From the perspective of IT management and control it is necessary to find out, which interrelations exist with shadow IT and what tasks are resultant. So far only little research exists on this topic. To overcome this deficit the presented project targets on a scientifically based definition of shadow IT, the investigation of best practices in several companies and the development and application of instruments for the identification, the assessment and controlling of shadow IT.**

Keywords- Shadow IT; IT Controlling; IT Governance; IT Service Management.

I. INTRODUCTION

IT management and control focus on the effective, efficient, transparent and compliant organization of information technology to achieve a best possible support of the business objectives [1]. This includes the minimization of risks and the recognition and realization of opportunities for improvements. The "official" IT infrastructure, developed, managed and controlled by the IT department, is supplemented in most companies by an unofficial IT. Business departments have a multiplicity of other hardware, software and IT employees. Generally these exist without the awareness, acceptance and support of the IT department. The resulting, autonomous developed systems, processes and organizational units are usually characterized as "Shadow IT" [2].

From IT management's perspective, some questions arise: What does the existence of shadow IT mean to its implementation? Does IT management have an influence on the growth or reduction of shadow IT? And what continuative tasks result from this subject?

Shadow IT is not a new phenomenon, but due to some current trends its significance is increasing [2]: New and primarily web-based technologies allow an easy access with low initial costs – so, on the first look, it is easy for a business department to select and get admirable IT services by itself. In addition to this the end users themselves play a particular role for growing shadow IT. Especially young employees have a strong bond to the usage of IT, as they grew up with it and use it in their daily private life. Thus, however, the expectation regarding the IT environment in their job is going to increase [3]. If the IT department is not able to satisfy their needs, the "emancipated" users start to take care of their IT devices and applications by themselves [4][5].

In this paper, we present first results of our research project "Shadow IT" [6]. Apart from the theoretical analysis of some detailed questions on this phenomenon and its definition, it is particularly necessary to develop methods for the identification and evaluation of shadow IT. In addition to that best practices have to be collected and the developed approaches have to be assessed in business. Several companies will be analyzed for data collection and for the verification of the methods mentioned above. All these steps are important to build a stable basis for developing an integrated and practical approach to control shadow IT. So far, we have set up the research concept and worked on the definition and the layout of the methods.

For this paper we will give a brief literature review in Section II. Based on this analysis, research questions are derived. In Section III we will present a detailed description of shadow IT and its occurrences. Section IV introduces the first concepts and developed methods for the identification and evaluation. Section V concludes with a brief outlook and next steps of the study.

II. STATE OF THE ART

This section examines the state of the art on the topic shadow IT. Therefore, an overview of the most considerable literature is given. Based on this, open research questions will be derived.

A. Literature Review

In spite of its rising significance, shadow IT has so far only attracted little attention in science. Some references can be found using the term shadow IT. But mostly, the topic has a tangential-role or it is only mentioned in connection with the main issue of the considered work. Most references are practical reports or blogs, which are based on the author's experience with no scientific foundation. Table 1 shows the central contributions, which are often referred to or which provide a solid investigation of the topic.

 ISBN: 978-1-61208-176-2

TABLE I. LITERATURE REVIEW

Reference	Main content
Sherman, 2004 [7]	This article focuses on Business Intelligence shadow IT, e.g., Excel- or Access-based systems, used to add information to reports, which are not supplied from the official IT. The systems start small and grow continually over time, which makes them costly to maintain. The data shadow systems can be recognized through user interviews on how reports are created. Sherman terms several reasons for their development: 1) Missing fulfillment of user's needs; 2) Shadow IT is easy to develop and seems to be "cost-free"; 3) A solution is needed, but the realization of official IT projects takes too long. To control shadow IT he suggests an improved communication between business and IT and the creation of data marts to secure consistent databases.
Bayan, 2004 [8]	Bayan describes reasons and effects of shadow IT and an approach how IT can deal with it. As the main reason he mentions the combination of reduced IT budgets and increasing IT demands. This forces business departments to develop their own IT. Furthermore, shadow IT is more focused on the business needs, it seems to be cheaper from business view and it appears to be faster and more dynamic than official IT. He refers primarily to security risks as the main effect of shadow IT. In his approach Bayan suggests to search for shadow IT with technical scanning tools. Afterwards, security gaps in the identified systems should be detected and closed. Finally, the implementation of new shadow IT should be reduced by achieving a better fulfillment of the business needs.
Jones et al., 2004 [9]; Behrens/ Sedera, 2004 [10]; Behrens, 2009 [11]	These publications refer to a study on a single shadow IT system in an Australian university [8]. The study describes an eight-year life cycle of a shadow software system, which was implemented and supported parallel to an official system. The work shows the possible reasons for its implementation and the opportunities and risks shadow IT can have. Furthermore, the work presents a few lessons learnt [10] on how management and the official IT should react on existing shadow IT. It is stressed that contrary to the common opinion shadow IT also has positive sides: It can be a source of innovation.
Raden, 2005 [12]	In his work Raden concentrates on spreadsheets for Business Intelligence. In his opinion this is the most common kind of shadow IT. These spreadsheets occur due to a lack of satisfaction of business requirements, such as reporting. Spreadsheets are an expressive, universally used, autonomous, fast and portable opportunity to fill these gaps. He highlights different problems through the behavior of developing shadow IT spreadsheets, e.g., wasted time, inconsistent business logic and inefficiencies. He concludes that a company-wide supply and integration of databases connected to all official IT systems can reduce the negative effects.
Schaffner, 2007 [13]	Schaffner describes effects and reasons for the development of shadow IT. As effects he lists several risks, such as poor engineering techniques, inefficiencies and compliance problems. His main argument for the existence of shadow IT is an insufficient alignment between business and IT. Typical efforts to reduce shadow IT, like the prohibition of shadow IT or the locking of administrator rights, don't have any effects. Schaffner suggests a closer cooperation between business and IT to increase the IT understanding of business processes and requirements.
Worthen, 2007 [14]	Worthen focuses on web tools and private devices as shadow IT. He highlights security and compliance violations as central risks. He is not considering the prohibition of shadow IT, because this could cause conflicts between business and IT departments. Also, the potential of user-driven innovations, which represents an opportunity of shadow IT, would be ignored. Instead he underlines, that the IT management has to find a strategy to deal with this subject. Worthen makes some general recommendations on how IT could handle shadow IT.
Shumarova/ Swatman 2008 [15]	In their study the authors focus on shadow collaboration systems, e.g., social software, wikis, etc. They explain the rising usage of these systems due to an easy and cost-free access and the growing merger of private and work life. They deduce and discuss three basic strategies on how to handle these shadow collaboration systems: 1) Rejection and banning; 2) Limitation and regulation; 3) Acceptance.
Dols, 2009 [16]	The topic of this master thesis is the search for causes of compliance defects in companies. In an empirical study, which analysis Dutch and Belgian subsidiaries of PwC, shadow IT is identified as one of two reasons for such defects. The work shows the state of the discussion on the topic shadow IT. Additional effects, causes or recommendations on shadow IT are not compiled.

B. *Open Research Questions*

The analysis of the articles listed in Table 1 and further existing references indicate a number of relevant open issues. These open research questions are listed in this paragraph.

1) Definition and theoretical framework: The term shadow IT is mostly described in an experience-related way. An academic, cohere and consistent definition of shadow IT and its classification according to a theoretical framework is missing.

2) Methods to deal with shadow IT: There are no specific methods or tools on how to deal with shadow IT. The existing frameworks and best practice approaches, such as ITIL [17] or COBIT [18], do not offer solutions regarding shadow IT. To develop a consistent methodology for this subject, the first steps are to identify shadow IT in practice and to evaluate the collected data. The developed methods need to be empirically tested. Best practices in the examined companies could be collected, to find out how successful companies deal with this topic. The answers to the first two research questions should establish a detailed basis for the following research work.

3) Business view: Most articles focus on shadow IT from IT view. The possibilities and consequences of the topic for the business are only analyzed occasionally.

4) Positive effects: The existing work primarily associates shadow IT only with negative effects. There is barely a focus on the opportunities of shadow IT. Nevertheless, to identify the potentials of user-driven shadow IT, it is necessary to identify positive outcomes of

 ISBN: 978-1-61208-176-2

shadow IT like improved process orientation and faster adoption of technical innovations.

5) Integrated approach: Mainly, the current contributions only focus on partial aspects of shadow IT. To handle the increasing phenomenon in practice and to give organizations an orientation on the controlling of shadow IT a balanced set of instruments and methods is necessary. Therefore, it is useful to collect best practices and develop an integrated, scientific approach including its relation to the different elements and tasks of IT management; IT governance and IT service management.

III. DEFINITION AND OCCURRENCES OF SHADOW IT

In Section I, we defined Shadow IT as a collection of systems developed by business departments without support of the official IT department.

This definition of shadow IT includes a variety of different occurrences [2]. One aspect is the usage of "Social Media Software" for business communication and data exchange or other services offered by providers from the internet, e.g., Cloud Computing or Software as a Service [14][19]. Furthermore, shadow IT includes the development and operation of self-built applications. In many cases these applications are Excel or Access based [7] and implemented by employees in the business departments. Moreover, the subject includes purchasing, in-house development and support of business intelligence solutions [12]. In the field of hardware, shadow IT relates to the integration of self-procured notebooks, servers, network routers, printers or other peripherals [13]. These devices are procured directly from a retailer, instead of being ordered via the official IT catalogue. A special case is the own purchasing of mobile devices, such as smartphones or tablets, and the usage of the related applications in the company network [20]. Finally, another occurrence is the development of own IT-support structures inside the business departments [12][13]: In case of IT incidents or problems technology-friendly colleagues are asked for help.

For the definition of shadow IT it is necessary to differentiate the term from end user computing (EUC). In this concept, the development of applications is delegated to the end users [21]. In contrast to shadow IT, EUC is officially initiated and supported. Primarily EUC is applied for the development of very easy IT solutions based on official platforms or for basic, individual configurations concerning specific applications.

The phenomenal description is one way to develop a definition for shadow IT. Another way is to consider existing work on informal organization [22] structures: Unofficial and hidden shadow IT processes are created in parallel with official structures. Similar to informal organization structures shadow IT differs from official policies and establishes own structures and processes. In addition, the emergence of both phenomena is linked with a distinct orientation towards employees' needs and results often from a lack within the formal structures, e.g., the autonomous acting of business departments pictures an irregularity concerning the decision of centralization within the defined IT governance.

Moreover, the emergence of shadow IT can be explained with information asymmetries and conflicts of interest between IT and business departments [23]. Information asymmetries associated with this relation exist as incorrectly understood business requirements by the IT and as a lack of knowledge by the business departments in general IT subjects and offered IT services. This asymmetry can lead to overpromised offers regarding service levels and software functionality and overcharged prices for IT services. The business departments experience these effects and therefore they try to reduce these risks. As a result, they deploy their own (shadow IT) solutions.

IV. IDENTIFICATION AND EVALUATION OF SHADOW IT

This section presents the current level of the research project in identifying and evaluating shadow IT. This refers to research question 2 and includes the collection of best practice data in the analyzed representative companies.

A. Identification Methods

Generally, there are three possible strategies for the collection of shadow IT information: 1) technical analyses [8]; 2) interpretation of help desk requests and 3) direct surveys of employees in the business departments [7].

The first approach is to identify shadow IT hardware or software with technical tools. Existing license management software and a network analysis tool for shadow hardware, which has been already developed in cooperation with this project team, can be used. The second method is based on information retrieved form the company's service desk. Incidents and problems identified there can be investigated on shadow IT as project experience proves that a remarkable number of calls is related to unofficial IT.

The third approach is a process oriented survey. It is based on structured interviews and process monitoring, to find out, which IT tools employees use in their daily business. Based on the experience gained in these interviews, we will try to develop standardized questionnaires to collect more information on user behavior and the usage of shadow IT.

The types of results from this identification phase are, e.g., graphical process descriptions with actual used IT tools and process-oriented IT landscapes with identified shadow

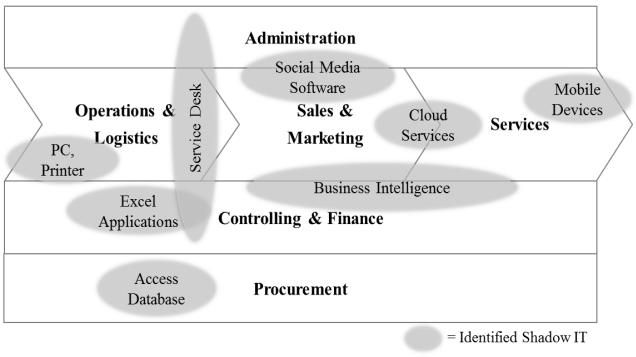


Figure 1. Process-oriented Shadow IT Landscape - Example

 ISBN: 978-1-61208-176-2

IT. Fig. 1 illustrates exemplarily the presentation of shadow IT in a process-oriented landscape on an abstract and high level. The identified shadow IT is assigned to one or several value chain activities [24], such as Operation or Administration. Also, the kind of shadow IT is shown. This type of representation can be refined on a level of departments and business processes to achieve a more detailed view. Thus, the process-oriented IT landscape allows picturing the shadow IT impact on business.

The different methods for shadow IT identification have certain advantages and disadvantages. The technical and help desk analysis enables a direct and quick search for shadow IT within the company's IT architecture. However, it is difficult to find all existing shadow IT occurrences with these techniques and it is not possible to define shadow IT related processes. In contrast to this, the structured interviews are based on the business processes and reveal the process-relation of identified shadow IT. However, this method depends on the knowledge and willingness of the interviewed users, e.g., the users might try to hide shadow IT applications from the interviewer. Furthermore, a lot of work and time is necessary to apply this method. Due to the described facts a combination of the methods is practical: The technical and the help desk analysis should be the foundation for the process survey. Thereby the expenses and disadvantages can be reduced and process-related results can be provided.

B. Evaluation Methods

After the identification of shadow IT, each specific system has to be evaluated. This validation is important to assess first needs of action due to risks. The evaluation results build the basic input for the development of guidelines and strategies. The following section briefly presents an evaluation model developed in the study.

For the evaluation, it is necessary to collect comprehensive information on the company and the IT, its policies and strategies. The general aim is to define aggregated characteristics to evaluate located shadow IT. Based on shadow IT examples in literature and discussions with companies and due to existing interactions of shadow IT with risk management, IT governance and IT service management topics, several parameters can be derived as mayor evaluation criteria.

The mayor criterion relevance describes the significance and importance of a located shadow IT instance for the investigated organization. Therefore, the analysis of the strategic relevance and the shadow IT criticality concerning the business processes, the IT security, the compliance and the IT service management is necessary. The mayor criterion quality refers to the system, the service and the information quality of the located shadow IT. Furthermore, the effects of shadow IT on the quality of business processing is of interest. The size of shadow IT is evaluated with regard to its use of resources and professionalism, its distribution in the company and its penetration with components and IT service processes.

TABLE II. SHADOW IT EVALUTAION CRITERIA

Shadow IT evaluation criteria		
Mayor criteria	***Sub-criteria level I***	***Sub-criteria level II***
Relevance	Strategic relevance	
	Criticality	Business process
		IT security
		Compliance
		IT service management
Quality	System quality	Hard-/Software
		Engineering process
	Service quality	
	Information quality	
	Quality of business processing	
Size	Use of resources and professionalism	
	Number of users	
	Shadow IT components	
	Shadow IT service processes	
Innovative potential		
Parallelism		

Apart from these criteria, it is essential to evaluate the innovative potential of the shadow IT instance. Finally, it is of interest to judge, if shadow IT is operated parallel to an existing, official IT-System or if it is complementary. Table II summarizes the different major and sub-criteria of this shadow IT evaluation model.

All sub-criteria on the different levels need to be weighted individually for the regarded company and rated for each located shadow IT instance. For the specific criteria evaluation different procedures and models, such as maturity models, can be applied. The total ratings of the major criteria are based on the weighted ratings of their sub-criteria. With these results each shadow IT instance is transferred into a

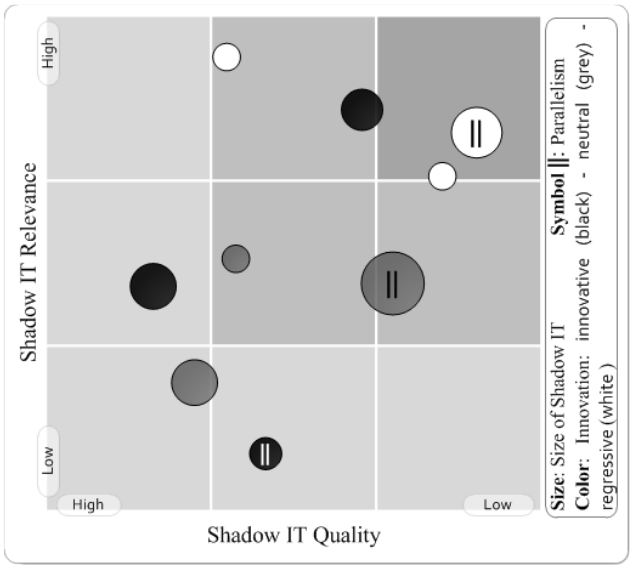


Figure 2. Shadow IT Evaluation Portfolio - Example

 ISBN: 978-1-61208-176-2

portfolio as exemplarily shown in Fig. 2. The portfolio consists of the two axes relevance and quality, the size for an instance and the color for the innovative potential. A parallel existing instance is marked with two parallel lines. The portfolio indicates which shadow IT instances have to be addressed with a high priority and establishes a basis for further management approaches to control shadow IT.

Furthermore, the development of shadow IT-related key performance indicators is an intended aim of the project. Based on this, it is possible to realize relevant benchmarks.

V. CONCLUSION AND FUTURE WORK

This paper introduced our research on shadow IT. The importance for IT management is shown and existing references are analyzed. As a result of this analysis, several open research questions could be pointed out. We have shown the initial steps and ideas of the research project with the focus on the definition, the identification and evaluation of shadow IT.

For the next steps, the theoretical questions on the definition of shadow IT and its relation to IT management disciplines have to be compiled. Besides, a detailed development of the discussed methods and their empirical appliance in practice will be carried out. Best practices for the handling of shadow IT will be investigated in the companies involved. Based on the results of data collection, the research project aims at the development of an integrated and practical approach to control shadow IT. This enables the revelation of its innovative potentials and the further development to a “User-driven IT”.

ACKNOWLEDGMENT

This research project is partially funded by the Ministry of Science, Research and Arts Baden-Württemberg [25]. The authors would also like to thank Cassini Consulting GmbH and Schutzwerk GmbH for supporting this project. Finally, we would like to thank the reviewers for their valuable input.

REFERENCES

[1] R. Zarnekow and W. Brenner, “Auf dem Weg zu einem produkt- und dienstleistungsorientierten IT-Management,“ HMD – Praxis der Wirtschaftsinformatik, vol. 40, no. 232, 2003, pp. 7-16.

[2] C. Rentrop, O. van Laak, and M. Mevius, “Schatten-IT: ein Thema für die Interne Revision,” Revisionspraxis – Journal für Revisoren, Wirtschaftsprüfer, IT-Sicherheits- und Datenschutzbeauftragte, April 2011, pp. 68-76.

[3] K. Quack, “Autoritätsverlust oder wahre Größe?” computerwoche.de, July 08, 2010, http://www.computerwoche.de/management/itstrategie/2349055/index.html, checked on 22/08/2011.

[4] Accenture GmbH, “Millennials vor den Toren – Anspruch der Internet-Generation an IT, ” Kronberg, 2009.

[5] RSA Security Inc., “The Confessions Survey,” Bedford, 2007.

[6] See for the research project “Shadow IT” our website www.schattenit.in.htwg-konstanz.de, checked on 22/08/2011.

[7] R. Sherman, “Shedding light on data shadow systems,” Information Management Online, April 29, 2004, http://www.information-management.com/news/1002617-1.html,checked on 22/08/2011.

[8] R. Bayan, “Shed light on shadow IT groups,” techrepublic.com, July 09, 2004 http://www.techrepublic.com article/ hed-light-on-shadow-it-groups/5247674, checked on 22/08/2011.

[9] D. Jones, S. Behrens, K. Jamieson, and E. Tansley, “The rise and fall of a shadow system: Lessons for enterprise system implementation,” ACIS 2004 Proceedings Paper 96.

[10] S. Behrens and W. Sedara, “Why Do Shadow Systems Exist after an ERP Implementation? Lessons from a Case Study”, PACIS 2004 Proceedings, Paper 136.

[11] S. Behrens, “Shadow Systems: The Good, the Bad and the Ugly,” Communications of the ACM, vol. 52, no. 2, 2009, pp. 124-129, DOI: 10.1145/1461928.1461960.

[12] N. Raden, “Shedding light on shadow IT: Is Excel running your business?” Hired Brains Inc., Santa Barbara, 2005.

[13] M. Schaffner, “IT needs to become more like "Shadow IT"," January 12, 2007 http://mikeschaffner.typepad.com/ michael_schaffner/2007/01/ we_need_more_sh.html, checked on 22/08/2011.

[14] B. Worthen, “User Management - Users who know too much and the CIOs who fear them,” CIO.com, 15/02/2007, http://www.cio.com/article/28821/User_Management_Users_Who_Know_Too_Much_and_the_CIOs_Who_Fear_Them_?page=1&taxonomyId=3119, checked on 22/08/2011.

[15] E. Shumarova and P. A. Swatman, “Informal eCollaboration channels: Shedding light on “Shadow CIT”,” eCollaboration: Overcoming Boundaries through Multi-Channel Interaction, 21st Bled eConference, June 15-18, 2008, Bled, pp. 371-394.

[16] T. Dols, “Influencing factors towards non-compliance in information systems,” UAS Utrecht, 2009.

[17] Office of Government Commerce, “ITIL - Service Strategy,” London: TSO, 2007.

[18] IT Governance Institute, “COBIT 4.1,” Rolling Meadows, 2007.

[19] B. Stone, “Firms fret as office e-mail jumps security walls,” International Herald Tribune January 11, 2007, http://www.nytimes.com/2007/01/11/technology/11iht-web.0111email.4167773.html?_r=1 checked on 22/08/2011.

[20] N. Zeitler, “iPad & Co. am Arbeitsplatz: Strategie gegen die Schatten-IT,” CIO.de, November 15, 2010, http://www.cio.de/_misc/article/printoverview/index.cfm?pid=157&pk=2250537&op=lst, checked on 22/08/2011.

[21] J.C. Brancheau and C. Brown, “The management of end-user computing: Status and Directions,” ACM Computing Surveys, vol. 25, no. 4, 1993, pp. 437–482.

[22] R. Lang: “Informelle Organisation,” in Handwörterbuch Unternehmensführung und Organisation, vol. IV, G. Schreyögg, A. von Werder, Eds. Stuttgart: Schäffer-Poeschel, 2004, pp. 497–505.

[23] V. Gurbaxani and C. F. Kemerer, Chris, “An Agent-Theoretic Perspective of the Management of Information Systems,” Proceedings of the Twenty-Second Hawaii Conference on Systems Science, vol. III, 1989, pp. 141-150, doi:10.1109/HICSS.1989.49234.

[24] M. Porter: “Competitive advantage: creating and sustaining superior performance,” The Free Press, New York, 1985.

[25] Ministry of Science, Research and Arts Baden-Württemberg, Germany – Website: http://mwk.baden-wuerttemberg.de, checked on 22/08/2011.

 ISBN: 978-1-61208-176-2

A Secure and Distributed Infrastructure for Health Record Access

Victoriano Giralt
Central ICT Services
University of Málaga
Málaga, Spain
e-mail: victoriano@uma.es

***Abstract*—The present paper describes the initial ideas for the author PhD thesis dissertation. The main goal of the research is to use Federated Identity and Access Management techniques in widespread use in academic networks, and more every day on the whole Internet, to the controlled, accountable and open access to health information over the Internet as well as controlling and securing the linkage of such data to a given individual. The challenge is to open the data buried in health records for research without giving out information that will allow to identify the individual persons. All of it keeping the real owners of the data, the individuals, in control of the information release. For this, we propose federated identity use to control access to linkage information about medical acts made publicly available. Using this technique, it would be even possible to provide totally anonymous informed health care.**

Keywords-health record; security; accountability; Federated Identity and Access Management.

I. INTRODUCTION

Health related data has the highest level of privacy protection in most countries data protection laws, but, at the same time, it is in the best interest of the whole medical science and the individuals themselves, that health data can be readily available.

The emergency room scenario has been used many times as an use case for expedited access to the whole heath record of an individual, where consent cannot be requested in the most life threatening situations. [1]

On the other hand, free access to high volumes of anonymous, but traceable (not to a real person only to an anonymous single individual), patient data, could be an invaluable resource for clinical research.

Access to health data should, in most cases, be granted by the individual to whom such data pertains, and should be accountable to those who see those data.

The present paper will propose a system than can be built using already available, and in use, protocols and tools that can both allow free access to anonymous health data and provide controlled and accountable means for de-anonymising the health records and tracing them back to the original person to whom they are related. [4][5][11][10] [12]

The proposed work builds upon the author experience in dealing with personal data in diverse scenarios, with some award winning results. [9] The driving force in the past eight years have been to put persons in the centre of their on-line lives and in control of personal data about them. [14][15] In this case, we propose a change of the status quo. At present, health records are owned by the institutions or practitioners that produce them, instead of the persons that are the subjects of those records. The main reason for our work is to put these persons (all of us) at centre stage and give them control over their own information, regardless of who has produced or created it. The present paper has resulted both from experience and the impression that the time is right for connecting two fields, health record management and electronic identity management, that are experiencing rapid development at this point in time. [13][11][1]

By publishing this work in progress paper, the author tries to gather as much hindsight as possible from others that might be working on ideas that could cross-pollinate and contribute to the final proposed landscape.

We will present the different scenarios of access and creation of health records by means of user stories:

- Individual enrolment
- Creation of health record in clinical practice
- Access to health records in clinical practice
- Access to health records from the emergency room
- Access to health information for research purposes
- Access to personal identity information

Finally, we will describe the technologies that will be used to create a demonstrator.

II. TECHNICAL TERMINOLOGY

The proposed work involves several domains with specialised terminologies that are not commonly understood. The author main field of work, despite his academic background, is electronic identity and privacy and access control, thus making this the main domain for the work.

A. Electronic identity terms

- Identifiable individual: A single physical person than can be identified by a set of personal data that constitutes their identity record.
- Attribute: A property of an identity record consisting of one or more values. All the values of an identity attribute are related by a common purpose or meaning. For example, the collection of telephone numbers belonging to

 ISBN: 978-1-61208-176-2

a person might form an identity attribute on the identity record that represents that individual.
- Principal: a person for whom another entity acts as an agent or representative.
- Pseudonym: an identifier that can single out an individual without revealing the real identity.
- Biometric information: personal information attributes derived from physical or biological characteristics of an individual.

III. General information processing and storage terms

- Hash: the result of using a hash function on an element of a data set. This functions transform larger data sets into smaller ones and produce the same result given the same input.
- Universally Unique Identifier (UUID): a 16 byte (128 bits) string that is guaranteed to be different from all other UUIDs generated before 3603 A.D., if the recommended algorithms are used [2].
- Resolver: an entity that can link pseudonymous identifiers like UUIDs to information about principals with or without identifying them.

A. Federated Identity and Access Management terms

- Identity Provider (IdP): An entity able to identify individuals and provide attributes pertaining to their identity.
- Relying Party (RP): An entity that trusts the federation and accepts identities asserted by IdPs.
- Federation: Infrastructure supporting the trust links between IdPs and RPs.
- Authorisation Server (AS): it is a trusted entity that takes access decisions based on attributes of the principals involved in a transaction in support of an RP.
- Attribute Authority (AA): is a trusted entity that asserts attributes about principals with or without revealing their identities to other principals involved in a transaction.
- Level of Assurance (LoA): the level of confidence with which the identity of an individual has been vetted in order to be linked to an electronic identity record.

B. Medical terms

- Health Level Seven (HL7): an international standards organisation that works for the interoperability of health clinical and administrative data. And, it is also used to refer to the standards defined by said organisation. [3]
- Act: one of the three main classes defined in the HL7 reference information model (RIM) [8] that represent actions that are executed and must be documented as various parties provide health care. [3]
- Role: second of the main classes defined in the HL7 RIM that establishes the function played by entities as they participate in health care acts. [3]
- Entity: third of the classes that represents the physical things and beings that are of interest to, and take part in, the health care. [3]
- Act Relationship: represents the binding of one act to another. [3]
- Participation: expresses an act's context, such as who performed it, for whom and where. [3]
- Role Link: represents relationships between individual roles. [3]
- Health Record (HR): a collection of health information related to an act or to the general health state of an individual. [3]

IV. Actors

A. Patient

We will use the term patient to refer to a person that is the subject of a medical act, although both in classical and modern medicine, keeping persons in a healthy condition is the main target of medical practice.

B. Practitioner

Practitioner will refer to any health care professional of any kind that interacts with patients in medical acts.

C. Emergency Room Practitioner

We have singled emergency room (ER) practitioners as they will receive special treatment in the system regarding the way they can access health records.

D. Staff member

This term refers to non medical professionals that have a role in medical acts like clinic receptionists or hospital administrative staff that require access to partial content of the HRs or to personal data of the patients.

E. Relative

A person with a family or other kind of social relationship to a patient that might play a role in authorising access to HR or provide personal information about the patient.

F. Researcher

A person that requires anonymous, or, at most, pseudonymous access to HRs for scientific research work.

V. General System Description

The proposed system aims to provide both freely available anonymous HRs published as HL7 [7] XML [16] documents on common web servers and a privacy controlled way of liking such records to the patients that participated in the corresponding medical acts.

There exist both commercial and non-profit repositories for personal health records, but they are centralised and, in many cases, under tight control of entities like insurance companies. We propose a totally open and distributed system based on trust models proven in higher education, research, government and vertical industries. The level of trust can be as high as to use one of such federations for controlling fusion nuclear reactors remotely or submitting experiments to synchrotron facilities.

 ISBN: 978-1-61208-176-2

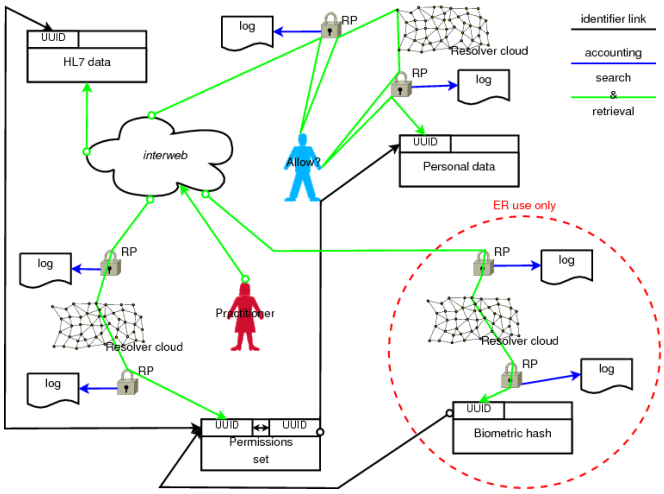


Figure 1. General system description

Identity federations are trust circles that have clearly defined rules for participation that, at the same time, act as codes of conduct for those entities that decide to participate in a certain federation. Even, in this moment in time, federations of federations, also known as inter-federations, are successfully being formed around the globe in the higher education and research sector. [11][10]

A. Pseudonymous identifiers

Any information part in the system, be it a patient personal information, an HR resulting from an act, or a biometric characteristic, will receive an UUID [2] as identification. These identifiers will be used as keys to find the resolvers that will link patients and HR using the UUIDs themselves as double indirection pointers. There will be a resolver finder cloud to locate the proper pointer to resolve a given UUID into the data that it represents.

B. HR publication

Web technologies facilitate the publication of huge amounts of data and also allow for easily indexing, locating and presenting such data. HL7 [7] XML [16] is a text format that provides all the required characteristics for easy web publication of HR, and, at the same time is an accepted interoperable format.

So, the information resulting from a medical act will be published as an HL7 XML document associated to an UUID that the practitioner or the relevant staff can provide to the patient in an electronic format. This UUID will be linked to the patient's UUID in a resolver the patient decides. This HR UUID to resolver relationship is also published through the resolver finder cloud.

The HL7 [7] document should not contain any personal information about the patient and the minimal possible amount of data.

C. Patient identification

The patients themselves will register to an IdP recognised in the global health care federation, using a method that provides an acceptable LoA. The personal data record will receive an UUID that can be published into the resolver cloud.

Patients should also get a hash out of some standardised biometric information. Ideally, this information should be genetic as it is the only type of biometric data that any body part carries. The state of the art does not yet allow for a full genomic characterisation of an individual in a reasonable time and for a reasonable cost, but there is fast progress in that area. Any other biometric information can be standardised, and, for the purpose of the demonstrator, we propose the use of digitised fingerprints, that will be hashed using Automated Fingerprint Identification Systems (AFIS) [17] algorithms.

Using fingerprints could be a handicap for the ER use case that we will present, in case the patient has lost the hands, but the prevalence of such situations is not high enough to render the system useless.

Once the patient has a biometric hash, it is associated to an UUID that will be published in a special resolver finder cloud, that allows for, so to say, backwards searches. This is required mainly for the ER use case.

The patient personal data will also include any relevant information needed for authorisation related contacts, either direct or through a relative.

D. Practitioner, staff and researcher identification

All other principals that participate in health care acts will register to pertinent IdPs in the federation, that could be run by hospitals, physicians or nurses colleges, insurance companies, etc. These IdPs will assert attributes that allow the AS, that control access to the RPs in the resolvers, to take appropriate decisions for granting access to the requested information. Thus, no one will get more information than that required to participate in a given act.

VI. User Stories

For the sake of brevity, we will do a shallow description of the user stories proposed in the introduction.

A. Individual enrolment

I'm a patient and want to publish my HR.

1) I select an IdP or the national health system provides me one.
2) I identify to the IdP using documents to achieve the required LoA and provide contact information for me and my closest relative.
3) I get the UUID that identifies my personal data.
4) My UUID is published by the IdP resolver.
5) My biometric hash is published in the resolver cloud.
6) I get my biometric hash UUID and link it to my UUID.

 ISBN: 978-1-61208-176-2

B. Creation of health record in clinical practice

1) I as a patient go visit a practitioner.
2) All acts are compiled into HR documents.
3) The HR are dated and get UUIDs.
4) The HR UUIDs and my UUID are inserted in my IdP resolver.
5) The HR UUIDs are sent to the resolver finder cloud from the resolver together with the pertinent pointer.

C. Access to health records in clinical practice

1) I as a patient go visit a practitioner.
2) The practitioner requests historic HR information.
3) I provide the practitioner with my UUID.
4) The practitioner identifies to the pertinent IdP and queries the resolver finder cloud and then, the appropriate resolver.
5) The resolver AS sends me a message indicating the practitioner identity, information about the requested data and a request for granting authorisation.
6) I grant the access and set a time limit.
7) The practitioner can access the data.

D. Access to health records from the emergency room

1) An unconscious and unidentified patient arrives in a life threatening condition.
2) The standard biometric parameters are determined and hashed appropriately.
3) A practitioner in the ER identifies to an IdP connected to an AA that asserts the attributes that verify the ER job.
4) The asserted attributes allow access to the special resolvers for biometric hashes, and to the UUID resolvers without requesting authorisation from the patient or relatives.
5) The resolvers return all HR UUIDs related to the UUID associated to the biometric hash.
6) The ER practitioner can retrieve the whole history of HRs related to the patient, without knowing the identity of the individual.

E. Access to health information for research purposes

1) I am a researcher working on a certain disease.
2) I search the web an collect all pertinent HRs.
3) I need to know about historic HR data about the same individuals that form the population under study.
4) I identify to my IdP that has an AA that asserts attributes to prove my researcher condition.
5) I query the resolvers for other HR UUIDs that belong to the same individuals as the HR UUIDs in the collection under study.
6) Depending on user preferences, data sensitivity or other parameters, patients get a request for granting access to the HR.

F. Access to personal identity information

1) I am a hospital staff member.
2) I need to know a patient identity for billing purposes.
3) I identify to the hospital IdP and the hospital AA asserts attributes to prove my administration staff status.
4) I query the resolver finder cloud to find the resolver for the patient UUID.
5) I query the patient resolver.
6) I get back the data needed to bill the patient.
7) The patient is notified of the personal data request.

VII. Technologies for Implementing the System

There are several options for some of the technologies needed to implement the proposed system. Producing a demonstrator is one of the main aims of the work described in the present paper, so it has been necessary to select a given technology for the different parts of the system. The selection has been mostly based on the author's experience or common practice in the fields in which he is working.

A. Security Assertion Markup Language (SAML)

SAML version 2 [4] is a proven method for expressing trust via electronic means and asserting information about principals that is in widespread use in present identity federations. It allow for inter-domain authentication, authorisation and accounting of access to resources. Such information is carried using XML [16] documents.

SAML2 will be used for the IdPs, AA and some RPs in the system.

B. Open Authorisation (OAuth)

Also in version 2, OAuth is a protocol that allows third party access to data with express authorisation of the owner of that data. [5]

OAuth will be used for the AS and some RPs in the system.

C. Distributed Hash Tables (DHT)

A distributed hash table (DHT) is a class of a decentralised distributed system that provides a look-up service similar to a hash table; (key, value) pairs are stored in a DHT, and any participating node can efficiently retrieve the value associated with a given key. Responsibility for maintaining the mapping from keys to values is distributed among the nodes, in such a way that a change in the set of participants causes a minimal amount of disruption. This allows a DHT to scale to extremely large numbers of nodes and to handle continual node arrivals, departures, and failures [6].

Due to the distributed, decentralised and resilient nature of DHTs, the system will use this technology to implement the resolver finder cloud. The keys will be UUIDs and the values will be URLs pointing to the resolver that can resolve a given UUID. In the special case of biometric hashes, the keys will be the later and the values will be UUIDs to feed into the normal resolver finder cloud.

 ISBN: 978-1-61208-176-2

VIII. A DEEPER VIEW OF USER STORY C

We will do a more detailed description of user story C, *Access to health records in clinical practice*, once we know the ttechnologieswe will be using for the ddemonstratorimplementation.

The actors and elements involved are:

- Patient: The subject in the clinical act.
- Practitioner: The health care professional performing the clinical act.
- IdP: The Identity Provider where the Practitioner authenticates.
- Resolver: The element that resolves the Patient identifier and locates pointers to HR.
- RP: The element that grants access to the Resolver.
- AS: The element inside the RP that permits the retrieval of pointers.
- HR: Relevant health information about the Patient.

Patient and Practitioner are both physical persons and their electronic representations, and computer applications acting in their name as proxies. The elements are computer applications and electronic representations of information.

Patient goes visit PPractitionerfor some clinical Act. Let's assume that is related to ccholesterolblood levels. It is the first time Patient and Practitioner meet. So, Practitioner needs some historic data about blood samples, mostly cholesterol levels and some related values. Thus, Patient provides Practitioner with a UUID that can be linked to published HRs, through the use of resolvers. The process flow proceeds as depicted in figure 2, with the following steps indicated as circled numbers:

1) Patient provides Practitioner with UUID
2) Practitioner goes to the resolver cloud
3) RP on resolver cloud requests Practitioner identity
4) Practitioner identifies to the pertinent IdP and returns to RP
5) AS in resolver cloud RP finds Patient authorisation method and requests access permissions for Practitioner.
6) The resolver AS sends Patient a message indicating Practitioner identity, information about the requested data and a request for granting authorisation.
7) Patient grants access and sets a time limit.
8) Practitioner retrieves a set of UUIDs from the resolvers that belong to previous Patient HRs with relevant information.
9) Practitioner retrieves the needed HRs.

RPs log all resolution requests and authorisation responses with pertinent identity information about the requestor and granter, in order to create audit trails.

In our ddemonstrator SAML2 [4] protocol will be used to carry identity and authentication information, while OAuth2 [5] will carry the authorisation requests and responses.

It is ppossibleto increase the security of the previous flow rrequiringPatient to also authenticate against an IdP for replying to the authorisation request.

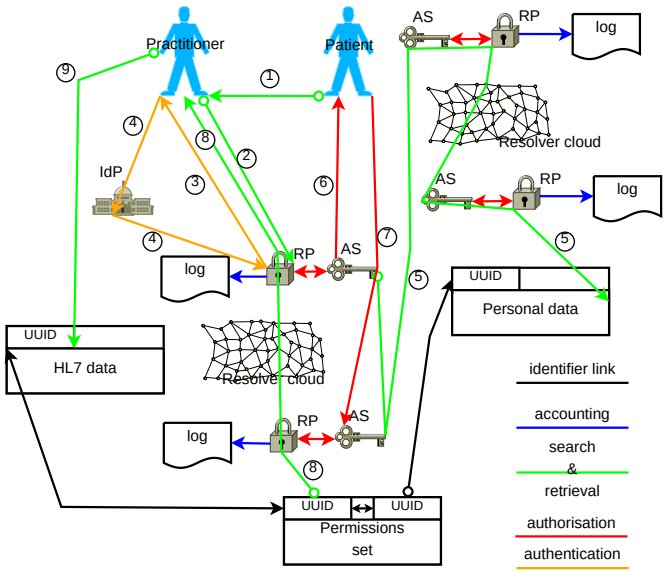


Figure 2. Access to health records in clinical practice

IX. CONCLUSIONS

In case the ideas presented in this paper are deemed worth the effort and such effort produces the expected results, the system will have two main advantages:

- a paradigm shift moving ownership of the data from the hands of those that produce such data into the hands of those to whom the data belongs to,
- and open data availability for many purposes.

ACKNOWLEDGEMENTS

The author wishes to thank all the fruitful conversations he has had with wise people in the Identity Federation space, with special mention of some ones that have seeded the ideas resulting in the work described in the present paper, including, but not restricted to, and in no particular order, Roland Hedberg, Ken Klingenstein, Andrew Cormack, J.A. Accino, Licia Florio, Klaas Wierenga, Milan Sova, RL "Bob" Morgan, Lorenzo Gil, Matthew Gardiner, Dave Birch, David Chadwick, and, last, but not least, for his very special support, Diego Lopez.

REFERENCES

[1] P. Groen, P. Mahootian and D. Goldstein, *Medical informatics: emerging technologies and 'open' health IT solutions for the 21st century*, January 2011, in press.

[2] ITU, *Universally unique identifiers*, URL: http://www.itu.int/ITU-T/asn1/uuid.html retrieved: November 21st, 2011.

[3] R. Gajanayake, R. Iannella and T. Sahama, *Sharing with care: An information accountability perspective*, Internet Computing, IEEE , vol.15, no.4, pp.31-38, July-Aug. 2011 doi: 10.1109/MIC.2011.51 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5749997&isnumber=5934 844 retrieved: November 21st, 2011.

[4] P. Madsen et al., *SAML V2.0 Executive Overview. OASIS Committee Draft*, April 2005. Document ID sstc-saml-tech-overview-2.0-cd-01-2col URL: http://www.oasis-open.org/committees/download.php/13525/sstc-saml-exec-overview- 2.0-cd-01-2col.pdf retrieved: November 21st, 2011.

 ISBN: 978-1-61208-176-2

[5] E. Hammer-Lahav, D. Recordon and D. Hardt, *The OAuth 2.0 Authorization Protocol*, http://tools.ietf.org/html/draft-ietf-oauth-v2-21 retrieved: November 21st, 2011.

[6] Wikipedia, *Distributed hash table*, http://en.wikipedia.org/wiki/Distributed_hash_table retrieved: November 21st, 2011.

[7] *Health Level Seven Standard Version 2.7 - An Application Protocol for Electronic Data Exchange in Healthcare Environments*, ANSI/HL7 V2.7-2011, National Institute of Standards and Technology, U.S. Government Printing Office: Washington, DC, 2011.

[8] *HL7 Version 3 Standard: Reference Information Model, Release 2*, ANSI/HL7 V3 RIM, R2-2010, National Institute of Standards and Technology, U.S. Government Printing Office: Washington, DC, 2010.

[9] V. Giralt, et al., *Example of Privacy Management in a Public Sector Organizational Electronic Directory* in Cunningham P., Cunningham M. (Eds.) Expanding the Knowledge Economy: Issues, Applications, Case Studies, IOS Press, Amsterdam, pp. 1386-1393, 2007

[10] V. Giralt, et al., *Federated Identity Infrastructure for the Andalusian Universities. Deployment of a Multi-technology Federation* in Cunningham P., Cunningham M. (Eds.) Collaboration and the Knowledge Economy: Issues, Applications, Case Studies, IOS Press, Amsterdam, pp. 1139-1144, 2008

[11] D. Simonsen, *Revealing the Identity of Federations*, the 16th European University Information Systems Organisation (EUNIS) congress, EUNIS 2010, University Information Systems: Selected Problems, University of Warsaw, Poland, pp. 11-20.

[12] M. Ramos, et al., *Design and Implementation Details of the Public Andalusian Universities Identity Federation CONFIA*, the 16th European University Information Systems Organisation (EUNIS) congress, EUNIS 2010, University Information Systems: Selected Problems, University of Warsaw, Poland, pp. 217-224

[13] *Microsoft® "Geneva" Server and Sun OpenSSO: Enabling Unprecedented Collaboration Across Heterogeneous IT Environments*, Microsoft® and Sun Microsystems White Paper, 2009, URL: http://download.microsoft.com/download/C/F/D/CFD1D9C8-EBA4-4780-B34B-DBEB5A4792B F/Geneva%20and%20Sun%20OpenSSO.pdf retrieved: November 21st, 2011.

[14] J.A. Accino, et al., *dUMA: comprehensive personal information management*, the 17th European University Information Systems Organisation (EUNIS) congress, EUNIS 2011, Maintaining a Sustainable Future for IT in Higher Education, Trinity College Dublin, Ireland

[15] J.A. Accino, M. Cebrian, and V. Giralt, *Identity Based Clusters of Applications for Collaboration and eLearning*, the 15th European University Information Systems Organisation (EUNIS) congress, EUNIS 2009, IT: Key of the European Space for Knowledge, University of Santiago de Compostela, Spain

[16] T. Bray, J. Paoli, C.M. Sperberg-McQueen, E. Maler and F. Yergeau eds., *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. W3C Recommendation 26 November 2008, URL: http://www.w3.org/TR/2008/REC-xml-20081126/ retrieved: November 21st, 2011.

[17] K.R. Moses, P Higgins, M. McCabe, S. Probhakar, S. Swann, *Fingerprint Sourcebook-Chapter 6: Automated Fingerprint Identification System (AFIS)*, National Institute of Justice/NCJRS 225326, 2010, URL: http://www.ncjrs.gov/pdffiles1/nij/225326.pdf retrieved: November 21st, 2011.

 ISBN: 978-1-61208-176-2

Active Mechanisms for Cloud Environments

Irina Astrova
Institute of Cybernetics
Tallinn University of Technology
Tallinn, Estonia
irina@cs.ioc.ee

Stella Gatziu Grivas, Marc Schaaf
Institute for Information Systems
University of Applied Sciences Northwestern Switzerland
Olten, Switzerland
{stella.gatziugrivas, marc.schaaf}@fhnw.ch

Arne Koschel

Faculty IV, Department for Computer Science
University of Applied Sciences and Arts Hannover
Hannover, Germany
arne.koschel@fh-hannover.de

Ilja Hellwich, Sven Kasten, Nedim Vaizovic, Christoph Wiens
Faculty IV, Department for Computer Science
University of Applied Sciences and Arts Hannover
Hannover, Germany
arne.koschel@fh-hannover.de

***Abstract*—Active mechanisms are used for the coordination (e.g., scalability) of IT resources in clouds. In this paper, we give an overview of existing technologies and products – viz., OM4SPACE Activity Service, RESERVOIR, Amazon SNS, IBM Tivoli Live Monitoring Service, Zimory and PESA – that can be used for providing active mechanisms in cloud environments. Our overview showed that these technologies and products mainly differ in the architectures they support and the cloud layers they provide.**

Keywords—Cloud computing; events; active mechanisms.

I. INTRODUCTION

Cloud computing has become more and more popular. Many companies (viz., cloud providers) are outsourcing their IT resources into clouds so that users can hire those resources only if they really need the resources and give the resources back when they do not need the resources any longer. This creates a new challenge for cloud providers – they need to provide users with systems, which can automatically assign IT resources on the fly. These systems should give the possibility to evaluate events from different event sources at one or more external coordination points. These points can coordinate the usage of the IT resources in clouds. Thus, the systems should use active mechanisms for the coordination (e.g., scalability) of IT resources in cloud environments.

The purpose of this paper is to give an overview of existing technologies and products that can be used for providing active mechanisms in cloud environments. Technologies like OM4SPACE Activity Service, RESERVOIR and PESA are mostly theoretical concepts and not end products. There are also (commercial) end products like Amazon SNS, IBM Tivoli Live Monitoring Service and Zimory.

II. OM4SPACE ACTIVITY SERVICE

OM4SPACE [2] provides software-as-a-service (SaaS), platform-as-a-service (PaaS) and infrastructure-as-a-service (IaaS). The crucial part of OM4SPACE is the Activity Service.

In cloud environments, often a large number of services occur on different layers. The Activity Service offers an approach for managing a large number of events from different event sources, processing these events and triggering appropriate actions on the events, e.g., starting new virtual machine instances when a specified threshold for the CPU load has been exceeded.

Figure 1 shows the architecture of the Activity Service, which consists of the following components:

- **Event Source:** This component can be an arbitrary part of a cloud environment; it generates different types of events (both simple and complex). Every event is sent to the Event Service for further processing. The Event Source can be on any layer of a cloud environment: SaaS, PaaS and IaaS.
- **Event Service:** This component receives events from an arbitrary number of Event Sources and performs the first step of processing, which is divided into two phases. The first phase dispatches the received events to Event Consumers that are registered for this type of events. The second phase consists of performing complex event detection (CED) on the incoming event stream. This CED can cause new complex and enriched events to be created by the Event Service and dispatched to the registered Event Consumers. Thus, the Event Service controls a granularity shift of the incoming event stream and helps to scale down the number of events, which enables complex event and rule processing by the Rule Execution Service.
- **Event Consumer:** This component receives a particular event type from the Event Service. Events can be of two types: simple events that the Event Service receives and complex events that the Event Service detects and generates. To receive events from the Event Service, an Event Consumer has to implement an appropriate event handler service, which needs to be published to the service registry.

 ISBN: 978-1-61208-176-2

The Event Service discovers event handler services by looking for the service registry. To inform the Event Service about the events an event handler service is interested in, filtering criteria have to be added to the WSDL description, which will be extracted by the Event Service.

- **Rule Execution Service:** This component receives events from the Event Service to match them against rules. Thus, it acts as an Event Consumer of the Event Service by registering an event handler service. Matching of the rules results in the execution of action handlers. An action handler needs to be implemented by each of the components that are to be called from within the rules. The rules are stored in the Rule Base, which is managed by the Rule Management Service.
- **Event Monitor (EM):** Not all components are built for active notification by the Event Service. For those components, a monitor capsule mechanism is used. As a result, a small application that acts as a monitor capsule around the Event Source can be implemented in such a way that it obtains events from the Event Source and transfers them to the Event Service. Furthermore, the monitor capsule can provide conversion between different types of events.

OM4SPACE adapts an activity service embedded in active database management systems (ADBMSs) to cloud environments. Active mechanisms are divided into different components that are put into the cloud. Each of these components has one or more well-defined interfaces. So the implementation of the components is interchangeable. The communication between the components is implemented using a service-oriented architecture (SOA). In addition to processing a large number of events, the Activity Service can be used for monitoring and scaling applications in the cloud.

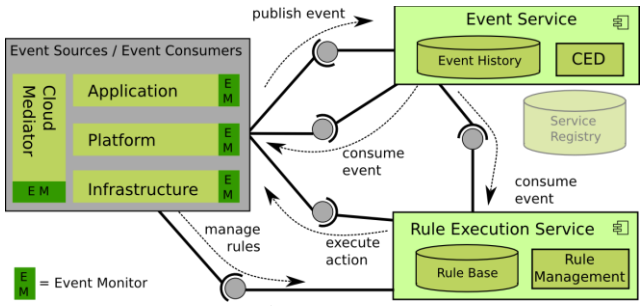


Figure 1. Architecture of OM4SPACE Activity Service [2].

III. RESERVOIR

RESERVOIR [3][4] provides IaaS. It requires the usage of manifests. A manifest serves as a contract between service and infrastructure providers.

Figure 2 shows the architecture of RESERVOIR, which consists of the following components:

- **Hypervisor:** This is a layer of abstraction, which runs on top of physical hardware. It allocates (physical) resources to virtual machines and manages and controls the execution of them by booting, suspending and shutting down resources as required. It can even provide the replication and migration of virtual machines. Examples of a Hypervisor include Xen [5] and VMWare [6].
- **Virtual Execution Environment Host (VEE Host):** This is the lowest layer in the architecture and provides plug-ins for different hypervisors. It enables upper layers to interact with heterogeneous virtualized products.
- **Virtual Execution Environment Manager (VEE Manager):** This layer implements the key abstractions needed for cloud computing and provides the functionality to control multiple VEE Hosts. Because of cross-site interactions between multiple different VEE Managers, the architecture offers the possibility to deal with and federate different sites, which implement heterogeneous virtualized products. Examples of a VEE Manager include OpenNebula [7].
- **Service Manager:** This layer is an interface to build the connection to the Service Providers. It ensures that the requirements of them are correctly enforced.
- **Service Provider:** This is the highest layer in the architecture and offers services to provide operations of specified businesses and uses the Service Managers to connect to the cloud.

RESERVOIR brings active mechanisms and the usage of events into cloud environments. The scalability of a service is enabled through an application description language (which introduces a monitoring framework along with Monitoring Agents) and key performance indicators (which describe the state of the service). The Monitoring Agents send Monitoring Events to the service management infrastructure, where these events are processed and rules are executed. The execution of these rules is additionally monitored by OCL operations to insure the correctness.

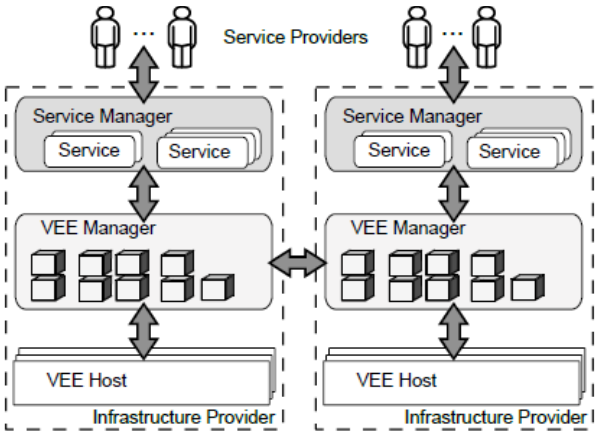


Figure 2. Architecture of RESERVOIR [3].

IV. AMAZON SNS

Amazon SNS (Simple Notification Service) [9] is a middleware product that offers a service for managing and sending notifications over cloud environments. The crucial part of Amazon SNS is topics.

 ISBN: 978-1-61208-176-2

Before sending notifications a topic has to be created. This is done by providing the topic name, which is used by Amazon SNS to generate a unique identifier called Amazon Resource Name. After the topic has been created, notifications can be sent to it. A notification consists of a message, the Amazon Resource Name and optionally a subject. Amazon SNS also adds meta-data like a signature and a timestamp to the notification.

To receive notifications, a subscriber needs to be registered for a topic. Every notification that arrives at a topic is delivered to all subscribers of this topic. Subscription contains endpoint information, which defines how the notification is delivered to the subscriber. Amazon provides the following types of endpoints:

- **Email:** The notification is sent via an email by using the SMTP protocol. The notification subject is mapped to the email subject and the notification message is mapped to the email message. Amazon SNS adds additional information to every outgoing notification, which contains an HTTP link for unsubscribing.
- **Email JSON:** The notification is also sent via an email but by using the human and machine-readable format called JSON [10]. All notification properties (e.g., the timestamp, the message and the subject) are stored in a list of key-value pairs.
- **HTTPS:** The notification is delivered by using the HTTPS protocol. In case of the incoming notification, Amazon SNS performs an HTTP-Post on a specified URL. The body of the HTTP-Post contains all notification information in the JSON format. All notification properties are encrypted and stored in a list of key-value pairs [11].
- **HTTP:** Like HTTPS but without using encryption.
- **Amazon SQS:** The notification is delivered to a queue of the Amazon SQS (Simple Queue Service). This is another Amazon service, which provides the functionality of sending text messages to the cloud. These messages are cached for a limited period of time while clients can request and receive them.

Every subscription needs to be confirmed by the receiver. For this purpose, Amazon SNS sends a confirmation request, which contains a specific token, to another endpoint. By transmitting this token back to Amazon SNS, the subscriber guarantees that it gets access to the endpoint and can receive notifications. Otherwise, it would be possible to enter arbitrary endpoints and flood them with notifications, which they do not want to receive. Amazon SNS supports an event-driven architecture (EDA) to decouple the notification sender from the receiver (i.e., the subscriber). Thereby it propagates this decoupling into the cloud.

Amazon allows for a detailed access control on Amazon SNS and its topics, e.g., by defining who is allowed to access a topic, who is allowed to add subscriptions to this topic and what types of subscriptions are permitted. For this purpose, Amazon introduces syntax for defining policies. These policies contain all information, which is important for security and access-control configuration.

Amazon SNS provides easy-to-use active mechanisms. But these mechanisms do not enable performing CED or condition-based rule execution. Another big problem with Amazon SNS is the small size of a message (8 kilobytes per notification).

V. ZIMORY

Zimory [13] provides IaaS. It is a product, where highly distributed components cooperate with each other using active mechanisms.

Monitoring and management are used to assure the scalability of virtual machine instances. The monitoring and management are done via a web interface through which users can specify rules that trigger one or more of the following actions when an event occurs [8]:

- **Storeback:** During this action, the virtual machine will be restored from a specified backup or set to a specified state.
- **Snapshot:** During this action, a snapshot of the current running state of the virtual machine will be created.
- **Clone:** During this action, the current running virtual machine will be duplicated (i.e., cloned). A clone can then be started immediately.

Zimory uses active mechanisms to distribute the CPU load to more than one virtual machine instance. In such a scenario, the cloud component (not named in the Zimory documentation), which monitors the instances, is acting like an event producing service that evaluates on what state the event should be produced. After this, the event is published to a component inside the cloud, which is able to receive events regarding the instances, a kind of event receiving service. This service then processes events and performs one or more of the actions listed above.

VI. IBM TIVOLI LIVE MONITORING SERVICE

IBM Tivoli Live provides SaaS. The crucial part of IBM Tivoli Live is the Monitoring Service [12].

The Monitoring Service is used to monitor network components and manage them in an online web portal. For this, a monitoring server is installed in the cloud. This server can monitor the cloud components (both active and passive) and send the collected data to the cloud. Here the data are stored and can be accessed via an online web portal. This portal gives users the possibility to set thresholds for the different monitored parameters of a component. When a threshold exceeds, an event is generated and an action for the event is performed. Unfortunately, the IBM documentation does not further specify how such an event or action can be used. But in theory it should be possible to send alarms via email to the users to notify them. Also, it should be possible to automatically perform an action on components, which use the Monitoring Service with an active agent. Such an action could be the execution of a script to automatically repair the state of a failed cloud component.

Figure 3 shows the architecture of the Monitoring Service. The following types of monitoring occur in the Monitoring Services:

 ISBN: 978-1-61208-176-2

- **Touchless monitoring:** A component, which is monitored, needs no further software installed, except a simple standard SNMP service. The Monitoring Service uses SNMP to retrieve current information about the component. This is only passive monitoring without any chance of interaction or controlling of the component. Also, the monitored data can be different across several components because SNMP does not standardize the monitored data but the messaging protocol.
- **Distributed monitoring:** This type of monitoring is based on an extra agent, which is deployed to the component being monitored. The monitoring server connects to this agent to retrieve information on the component. With the agent solution, it is also possible to control and manage the component.
- **Performance services:** These services are used to monitor data for long-time performance analysis and bottleneck indication. It should be noted that the performance services are used for manual analysis only.

The Monitoring Service uses multi-agent systems for active mechanisms. So it is more aligned to the "old-fashioned" IT systems. But the Monitoring Service can also be used in cloud environments because such environments are nothing else than virtualized IT systems.

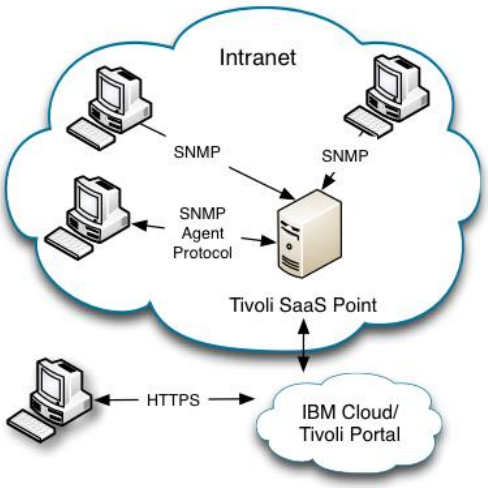


Figure 3. Architecture of the IBM Tivoli Live Monitoring Service [12].

VII. PESA

PESA [11] is a technology that adds policies to an event-driven service-oriented architecture (ESA). A policy is a set of rules that control behaviors of services.

A large number of services that are deployed into different cloud environments can generate a large number of events. Because of that, it is not practical to set up a centralized management service that collects, stores, correlates and processes events. A better approach would be to build management services responsible for a small number of services, thereby reducing bottlenecks and speeding up the reaction of the whole system. PESA offers this approach. The management services are used for:

- Matching a pattern that consists of different single events, which can indicate some failure.
- Creating high-level business events out of simple events, which can indicate that some reaction of the business layer is needed to assure the execution of the business workflow.
- Adding data to events to give more information about the reason for the occurrence of an event.
- Triggering conditions for policies or rules to react on occurrence of events without human interaction.
- Invoking services or business workflows when their conditions are satisfied to assure that a business workflow can be executed.

As the services they are managing, the management services should be loosely coupled and highly distributed so that they can be set up in different cloud environments near the managed services. This will improve the performance while managing the services, reduces the complexity and eliminates drawbacks of a centralized management service (e.g., single points of failure). In such a scenario, the managed services can automatically incorporate service components for monitoring and management from their "local" management services. Furthermore, it becomes possible to compose the "small" management services into a "bigger" management service that provides coherent management for all services used in the business workflow.

Figure 4 shows the architecture of the management service. The components of the architecture can be assigned to one of the following layers:

- **Monitors and sensors, extraction and transformation tools:** Events are generated through monitoring services and sensors. Then every occurring event is transformed to a standard data format by a tool and the resulting events are published to an enterprise service bus (ESB). Both layers correspond to the event generation layer of an EDA.
- **Classification and categorization:** This is the layer where events are classified and assigned to a class of events in order to provide a coherent scope on the events for the layer above.
- **Analysis engines:** This layer correlates, associates and links events together to generate more complex events that have relevance to business. It corresponds to the event processing layer of an EDA.
- **Operational actions, policy actions and conflict resolution actions:** These are the layers where policies trigger actions based on events as their conditions. So the layers are responsible for performing the actions that assure the execution of a business workflow. The layers correspond to the event handling layer of an EDA.

All the layers communicate with each other through the ESB. Layers are built so that higher layers have broader scope that enables more complex analysis and management.

 ISBN: 978-1-61208-176-2

But it does not mean that the high and low-level management services are built with more or less complex components. Rather every management service is able to handle events and perform actions based on policies. Such a system is highly extensible, by adding more management services responsible for managing more services.

The intelligence that PESA adds to active mechanisms is the possibility to build business workflows that are very agile and can be easily adapted to changes, without human interaction during the execution of a business workflow. This becomes possible only by adding the events to signal changes and by adding the policies to react on those changes.

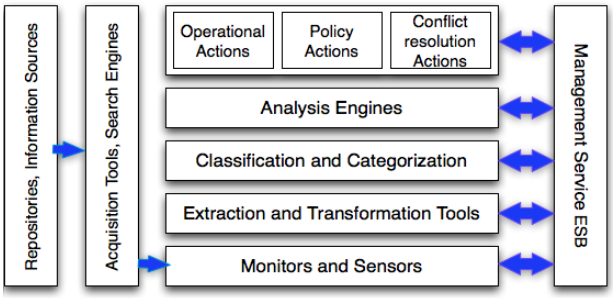


Figure 4. Architecture of the PESA management service [11].

VIII. CONCLUSION

This paper gave an overview of existing technologies and products – viz., OM4SPACE Activity Service, RESERVOIR, Amazon SNS, IBM Tivoli Live Monitoring Service, Zimory and PESA – that can be used for providing active mechanisms in cloud environments. Table I summarizes this overview.

TABLE I. SUMMARY OF OVERVIEW OF TECHNOLOGIES AND PRODUCTS FOR PROVIDING ACTIVE MECHANISMS IN CLOUD ENVIRONMENTS

Technology / Product	**Architecture**	**Cloud layer**
OM4SPACE Activity Service	SOA	SaaS, PaaS, IaaS
RESERVOIR	SOA	IaaS
Amazon SNS	EDA	PaaS
IBM Tivoli Live Monitoring Service	SOA	SaaS
Zimory	EDA	IaaS
PESA	ESA	PaaS

The overviewed technologies and products mainly differ in the architectures they support: EDA, SOA or ESA.

An EDA [1] enables event producers to publish their events and event consumers to subscribe to and consume those events. In an EDA, events know nothing about their consumers. Events can also remain unconsumed because none of the consumers is interested in them. There are no direct relationships between the event producers and consumers. So the services built on top of an EDA are loosely coupled. This helps an EDA fit into a scenario where the services deployed into the cloud are managed by management services. Support of an EDA can be found in Amazon SNS and Zimory.

A SOA enables the composition of loosely coupled highly distributed services. These services can be deployed into different cloud environments where the clouds themselves take care of the services. Support of a SOA can be found in OM4SPACE Activity Service, RESERVOIR and IBM Tivoli Live Monitoring Service. The Activity Service in its initial version follows a cloud-native approach, by using an ESB to build a SOA and Web services to provide communication between loosely coupled highly distributed components. RESERVOIR also supports a SOA but it does not actually specify such communication. The Monitoring Service follows a more old-fashioned approach, by using multi-agent systems for active mechanisms. One possible reason for this is the growing structure of the whole IBM Tivoli Live.

An ESA [11] is the result of combining an EDA with a SOA. Such a combination is needed because a SOA typically composes services to business workflows. It does not account for events that occur across or outside of business workflows or complex events. Being combined with an EDA, a SOA can react on events. For example, a high-level business event can cause the execution of a single service or a set of services that can handle the problem occurred in a business workflow. Such a SOA enriched by events through an EDA can be used to build agile business workflows that adapt to changes, which occur during the execution of the business workflow. In such a scenario, the changes will be signaled by events. An EDA can also take advantage of the combination with a SOA because of the flexibility that a SOA provides through composing of services on different layers. As a result, it becomes possible to integrate an EDA on every layer. So an EDA can become responsible for publishing, subscribing and consuming events on both simple low service layer and high complex business layer. Because of these advantages, an ESA fits well into a scenario where events should be monitored, enriched and connected with each other on different levels. The connection between events is important because it can be used to connect multiple low-level system events to create a high-level business event. In such a scenario, events can occur everywhere, e.g., they can be created from applications, databases or services that are involved in a business workflow. Support of an ESA can be found in PESA. One possible reason for this is that cloud environments are typically environments for loosely coupled highly distributed services that can be orchestrated to a business workflow.

The overviewed technologies and products also differ in the cloud layers on which they provide their services: SaaS, PaaS or IaaS.

SaaS is a model of software deployment whereby a cloud provider licenses an application (i.e., software) to users for the usage as a service on demand. OM4SPACE Activity Service and IBM Tivoli Live Monitoring Service provide SaaS. One possible reason for this is the popularity of SaaS. (Currently, SaaS is the most popular type of cloud computing because of its simplicity, flexibility and scalability.)

 ISBN: 978-1-61208-176-2

PaaS is a model of application development and delivery. In particular, PaaS offers a development platform for users. OM4SPACE Activity Service, Amazon SNS and PESA provide PaaS.

Whereas SaaS allows for the usage of applications in cloud environments and PaaS offers the ability to develop and deliver these applications, IaaS provides users with the infrastructure for developing, running and storing the applications. RESERVOIR and Zimory provide IaaS. OM4SPACE Activity Service could also be used as an event processing component within IaaS.

IX. FUTURE WORK

There are advantages and disadvantages with all the overviewed technologies and products. An advantage of one is often a disadvantage of another. Therefore, the most promising approach would be to combine them all. This approach could be based on OM4SPACE Activity Service because it could possibly operate or be utilized on all different layers of a cloud environment and performs the complete event roundtrip by: generating events at an Event Source, sending events to the Event Service, performing CED at the Event Service and generating new complex events, sending events to an Event Consumer and the Rule Execution Service, performing rule processing and rule execution, and performing action invocations on action handlers in case of matching rules.

The Activity Service could benefit from using RESERVOIR. RESERVOIR defines a standard way to monitor a cloud component in order to read the parameters out of that component using a manifest, which assures that the Service Providers can deploy their services into the cloud. The Activity Service should also be able to monitor cloud components (both active and passive), but without concretely defining what parameters would be monitored. This may first not be seen as a real problem. But when rules for the events should be generated, it may become a big issue because the rules use the attributes of the events, which are set at a cloud component. Also, for the content enrichment of events at the Event Service, a concrete set of attributes should be defined. Thus, the usage of RESERVOIR could solve the problem of defining rules and getting the parameter dependencies for the rules.

Another improvement could be done in the action performing part of the Activity Service, which is currently implemented as a simple call to an action handler. This could be improved if the action handler and the Rule Execution Service used Amazon SNS as a transport mechanism. For example, the action handler could subscribe to a topic filled by the Rule Execution Service via Amazon SNS. But this problem cannot be solved by using Amazon SNS solely because this product is mostly not standardized and thus, can cause a vendor lock-in. Another problem with the usage of Amazon SNS is the small message size (8 kilobytes only). Therefore, a better solution would be if the Activity Service itself could implement a-la Amazon SNS transport mechanism. Such independence from a transport mechanism is currently implemented in the Activity Service to allow for better integration with existing cloud communication services.

ACKNOWLEDGMENT

Irina Astrova's work was supported by the Estonian Centre of Excellence in Computer Science (EXCS) funded mainly by the European Regional Development Fund (ERDF).

REFERENCES

[1] J. Dunkel and R. Bruns. Event-Driven Architecture. Springer, 2010.

[2] A. Koschel, M. Schaaf, S. Gatziu Grivas, and I. Astrova. An Active DBMS Style Activity Service for Cloud Environments. 1st Intl. Conf. Cloud Computing 2010, pages 80–85, IARIA, Portugal, November 2010.

[3] C. Chapman,W. Emmerich, F. G.M´arquez, S. Clayman, and A. Galis. Software architecture definition for on-demand cloud provisioning. 19th ACM International Symposium on High Performance Distributed Computing, HPDC'10, pages 61–72, New York, NY, USA, 2010. ACM.

[4] S. Eliot. Reservoir homepage. http://www.reservoir-fp7.eu/ Accessed: November 2011.

[5] Citrix Systems Inc. Xen. http://www.xen.org/ Accessed: November 2011.

[6] VMware Inc. Vmware. http://www.vmware.com/ Accessed: November 2011.

[7] OpenNebula. Opennebula. http://www.opennebula.org/ Accessed: November 2011.

[8] F. Galan, A. Sampaio, L. Rodero-Merino, I. Loy, V. Gil, and L. M. Vaquero. Service specification in cloud environments based on extensions to open standards. 4th Intl. ICST Conference on Communication System software and middleware, COMSWARE'09, pages 19:1–19:12, New York, NY, USA, 2009. ACM.

[9] Amazon Simple Notification Service (Amazon SNS). http://aws.amazon.com/de/sns/ Accessed: November 2011.

[10] Json (javascript object notation). http://www.json.org/ Accessed: November 2011.

[11] P. Goyal and R. Mikkilineni. Policy-based event-driven services-oriented architecture for cloud services operation and management. IEEE Intl. Conference on Cloud Computing, pages 135–138, 2009. IEEE.

[12] IBM Tivoli foundations and IBM Tivoli Live Monitoring Services. http://www-01.ibm.com/software/tivoli/products/monitor/ Accessed: November 2011.

[13] Z. GmbH. Zimory Enterprise Cloud Anwendungsbeispiel. http://www.zimory.de/index.php?id=75 Accessed: November 2011.

 ISBN: 978-1-61208-176-2

Information Technology Planning For Collaborative Product Development Through Fuzzy QFD

Jbid Arsenyan
Industrial Engineering Department
Bahcesehir University
Istanbul, 34100, Turkey
jbid.arsenyan@bahcesehir.edu.tr

Gülçin Büyüközkan
Industrial Engineering Department
Galatasaray University
Istanbul, 34357, Turkey
gulcin.buyukozkan@gmail.com

***Abstract*—Collaborative Product Development (CPD) becomes a more complex process to manage by the rapid technological change. As a consequence of various system features introduced by research groups and commercial packages, CPD practitioners lose track of the available platforms, protocols, applications, system features, and tools supporting CPD processes. This study aims to provide a mapping between the technological requirements for CPD and system features of these various infrastructures. Fuzzy Quality Function Deployment (QFD) is employed for mapping between requirements and features. An industrial expert is consulted for evaluation of derived relationships and consequently system features are prioritized.**

Keywords: Collaborative Product Development; Fuzzy QFD; Technology requirements.

I. INTRODUCTION

Due to its technology-centric nature, Collaborative Product Development (CPD) is typically based on technological infrastructures, which require for information technologies (IT) to be essential conveyors of good CPD performance [1]. However, the management of requirements and implementation of necessary tools to respond to these requirements constitute a complex process as the technological diversity grows rapidly. Current tools become hard to track and thus, evaluations are performed with incomplete and biased information given that assessing all systems is not possible.

Previous studies do not propose a comprehensive review of CPD systems mainly because these systems including various applications, tools, and plug-ins are numerous; they can be easily outdated by new researches and are only known by a limited community. On the other hand, various systems are proposed by literature and commercial ventures in order to facilitate collaboration, integration, co-design, and co-development processes of CPD teams.

In this highly uncertain environment, with various different requirements and numerous technological solutions, a systematic methodology is essential to plan the technological infrastructure needed to start and maintain the CPD process. Determining requirements and accordingly prioritizing technological response compose an important phase in IT planning. Some projects may require only communication tools, while others are dependent on highly skilled web-based engineering applications. A comprehensive and detailed planning methodology utilizing Quality Function Deployment (QFD) is introduced to help CPD practitioners in their development and collaboration efforts.

QFD is a well established methodology in transforming customer needs into engineering characteristics and therefore its House of Quality (HoQ) diagram appears to be a suitable tool for mapping needs of CPD into existing tools and technologies. Additionally, mapping is performed under a fuzzy environment in order to translate linguistic evaluations of the expert into quantifiable performance measures. The aim of this study is to introduce a comprehensible methodology for IT planning, which can be employed by CPD practitioners before launching CPD projects.

The study is organized as follows: next section introduces the technology planning literature, which covers studies in a general context. Then the fuzzy QFD is described and the methodology backgrounds are established. The fourth section presents CPD technology overview, which includes commonly used standards and environments, technology requirements and system features in CPD infrastructure. Then the IT planning with fuzzy QFD is presented with an evaluation of an industrial expert. The study concludes with a few remarks.

II. TECHNOLOGY PLANNING BACKGROUND

Use of proper technology is the most preferred factor in maintaining competitive advantage [2]. Systematic planning of technological infrastructure is therefore important in improving CPD performance. Efficiency and effectiveness of CPD are enhanced by appropriate implementation of tools and technologies enabling CPD [3], which can be attained through accurate mapping of requirements into the system features.

A technology planning framework is proposed by Porter et al. [4], which includes technology forecasting, as well as environmental analysis and aims to design organizational actions. Value adding chain concept requires the implementation of technology within all aspects of the business. Martin [5] also starts with technology forecasting

 ISBN: 978-1-61208-176-2

and applies scenario analysis to define technology allocations according to short term and long term needs. Rip and Camp [6] propose a four step methodology, which starts with market research then determine product features and technology options for these features and finally finishes with future consideration of technology resources.

Pretorius and Wet [7] define a framework based on the hierarchy of the enterprise, business processes and functions. Technological assessment can be mapped on the relationship between technology and processes on the three dimensional framework. Kumar and Midha [8] employ. They utilize the QFD approach to compare company's requirements in CPD with different functionalities of Product Data Management (PDM) systems and technical specifications are then compared to a specific PDM system.

Büyüközkan et al. [3] present a comprehensive review on tools, techniques and technologies enabling agile manufacturing in concurrent Product Development (PD). Rodriguez and Al-Ashaab [9] identify CPD supporting system characteristics and classify corresponding technological requirements. They also perform a survey in injection mould industry and they propose a knowledge based CPD system architecture responding to industrial requirements.

Koc and Mutu [2] present a technology planning methodology, from selection of competitive priorities to designing the activities, by integrating different system design perspectives through AD. Rueda and Kocaoglu [10] state that market and technology performance uncertainty make technological investment highly risky and they focus on diffusion of emerging technologies. They combine bibliometrics analysis, Delphi method, utility curves, and scenarios to define a composite indicator for the diffusion. Shengbin et al. [11] focus on technology roadmap concept and they present a visual guide to map market, product, and technologies to achieve technology selection. The three phased design process includes trend discussion, industrial and academic investigation, expert feedback on technological demand and it provides a tool to make strategic level technology selection decision.

Luh et al. [12] combine Design Structure Matrix with Fuzzy Sets Theory into FDSM to present a dynamic planning method for PD, increasing PD efficiency and decreasing development time. Ko [13] also employs FDSM to present a methodology enhancing PD management by organizing design activities and measuring dependency strength. Palacio et al. [14] presents a tool to facilitate collaboration in distributed Software Development (SD) teams, which aims to increase collaboration awareness by focusing on individuals and their activities.

Previous studies do not address a generic approach, which investigates and classifies the CPD requirements, as well as the tools and techniques provided by researchers and commercial packages. This study aims to introduce a planning framework within the fuzzy HoQ in order to capture these aspects and map their relationships.

III. FUZZY QFD OVERVIEW

HoQ, the planning tool within QFD methodology, can be described as a "conceptual map that provides the means of inter-functional planning and communications" [15]. It translates customer needs into customer attributes (CAs) in order to meet them through engineering characteristics (ECs).

As a first step in constructing a HoQ, CAs are collected from customers (Domain 1). Then engineering teams try to answer the question "how to achieve this attribute". ECs that affect CAs are listed accordingly (Domain 2). CAs are prioritized in order to have a trade off basis in the case of conflicting objectives (Domain 3). As depicted in Fig. 1, right hand side of HoQ offers a benchmarking tool, where customer perception of other brands as well as focal firm's brand in response to CAs is depicted (Domain 3a).

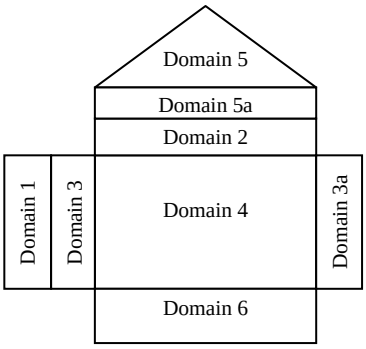


Figure 1 Main domains of HoQ

Then relationships between CAs and ECs are represented in symbols in accordance with the strength of the relationship (strong positive, medium negative, etc.). This step of the methodology serves to identify how an EC can affect a specific CA (Domain 4).

ECs effect on each other is represented in the roof matrix of the HoQ (Domain 5). Interdependent characteristics are thus displayed and the total outcome of engineering change is visualized. ECs are also marked regarding the direction of change in that specific characteristic (Domain 5a). Finally, target values and the degree of technical difficulty are set for ECs in order to present the amount of work and its complexity (Domain 6).

Majority of QFD applications stop at the planning stage, i.e. the HoQ and nevertheless, many benefits can be achieved through only the first matrix [16]. However, conventional HoQ matrix is not sufficient in describing the relationships between CAs and ECs. In some cases, application is performed in a fuzzy environment. Fuzzy QFD is employed in these cases in order to translate the vagueness of relationships and the subjectivity of the evaluator into quantifiable data.

Literature proposes many examples of fuzzy QFD applications. Şen and Baraçlı [16] investigate enterprise software selection requirements with fuzzy QFD. Linguistic variables are employed to prioritize non-functional criteria

 ISBN: 978-1-61208-176-2

in order to provide a decision making framework to determine the order of criteria to be satisfied during software selection decisions of a company. In their two concurrent studies [17] and [18], Vinodh and Chintha investigate the enabling effect of fuzzy QFD to leanness and agility in a manufacturing organization. Fuzzy QFD is employed to prioritize the lean competitive bases, lean attributes, lean enablers in one case and the agile decision domains, agile attributes and agile enablers in the other by employing linguistic terms for both relationship matrix and correlations.

Lee and Lin [19] employ fuzzy QFD in PD. They incorporate fuzzy Delphi, fuzzy Interpretive Structure Modelling and fuzzy Analytic Network Process into QFD framework. Linguistic variables are employed for both relationships between CAs and ECs and the correlation between CAs to investigate priorities of PD in CAs, ECs, part characteristics, key process operations, and production requirements. Liu [20] employs fuzzy QFD to investigate priorities in product design and selection by (1) computing the relative importance of CAs, (2) computing the final importance of CAs and (3) computing the final importance of ECs through linguistic variables. Their methodology is also two phased, the second phase adopting a multi-criteria decision making approach. Jia and Bai [21] apply fuzzy QFD in manufacturing strategy development. Fuzzy integrated HoQ helps to capture the highly imprecise and vague nature of the strategy decisions.

In this study QFD is employed in a fuzzy environment considering that IT planning of CPD projects, in terms of requirements and features, is dependent on subjective judgments of CPD managers,. We aim to translate subjective and linguistic judgments of evaluators into quantifiable relationships by integrating fuzzy sets theory into HoQ. In the proposed methodology; CA weightings, CA-EC relationships, and EC correlations are defined in linguistic terms and then translated into triangular fuzzy numbers (TFNs) in the form of (l, m, u). After defining CAs and ECs for the study, the industrial expert is consulted for his judgments. Collected linguistic judgments are fuzzified.

Fuzzy computation processes in this study are adapted from Vinodh and Chintha [18]. The relationship matrix and the weights of CAs are employed to compute the relative importance of ECs as follows:

$$RI_j = \sum_{i=1}^{n} W_i \otimes R_{ij}, j = 1, \ldots, m. \quad (1)$$

Then the correlation matrix is considered. The final score of the j^{th} EC is computed by the following equation:

$$score_j = RI_j \oplus \sum_{j' \neq j} T_{jj'} \otimes RI_{j'}, j = 1, \ldots, m. \quad (2)$$

The final score is defuzzified in order to obtain a final crisp score:

$$S_j = (l + 2m + u)/4 \quad (3)$$

The ECs are ranked in decreasing order of crisp scores. A higher score of EC implicates a higher priority to consider and thus, a higher importance to attribute.

IV. CLASSIFICATION OF IT FOR CPD

Technological change, especially in PD and collaborative technologies domains, are increasingly rapid and hard to track. However, services offered by various systems do not transform in the same pace as the complexity level increases.

CPD systems are generally built on various infrastructures. Commercial software and academic projects based on these infrastructures are numerous to cite and easily outdated, therefore out of the scope of this research. Nevertheless, some systems and commercial packages, summarized in [9] and [22], can be a reference on the services offered by researchers and industry.

A. Requirements overview

CPD literature and industrial experts express similar opinions when it comes to technological requirements in CPD projects, although some differences may be observed. Li and Su [23] state that CPD environment should comprise *scalability*, *openness*, *heterogeneity*, *resources access and inter-operation*, *legacy codes reusability*, *and artificial intelligence* as features. According to Rodriguez and Al-Ashaab [9], common access of design information, collaborative visualization of the component, and collaborative design of the component are the requirements to be supported by collaborative technologies. Palacio et al. [14] classify SD requirements in four groups: *scale*, *uncertainty*, *interdependence*, *and communication.* These requirements form a starting point for both collaboration and development processes. Requirements of CSD, which can be viewed as CPD sub-domain, include *interaction*, *knowledge*, *awareness*, *coordination*, *communication*, and *control* [14].

These aspects are categorized in nine groups under the *Requirements* domain, each requirement followed by its label.

Communication (CA1) emerges as a principal requirement in IT planning to assure awareness [22]. *Project Management* (CA2) and *Knowledge Management* (CA3) are two essential requirements as stated in [1,3,9,24], which clearly suggest that these two requirements should be considered within any type of project, regardless of its collaborative aspect. Another important requirement while planning the technological infrastructure of CPD is the *product model* (CA4) itself. The technological infrastructure should comprise a system that enables the representation, visualization, modification of the product model. *Data Integration & Analysis* (CA5) requirement can be described as a mechanism to integrate data available on different sites from different collaborating teams and to analyze this data in a most efficient manner [25]. Accordingly,

Interoperability (CA6) requirement emerges as a natural result of collaboration in order to assure diverse systems to work together.

Security (CA7) and privacy issues arise as CPD projects become a part of the business routine. This requirement implicates data protection as well as system back-up, as mentioned in [1]. Accordingly, defined by ISO 31000 as the effect of uncertainty on objectives, *Risk Management* (CA8) is a requirement to control uncertainties that may result in project failures. Lastly, CPD infrastructure requires *Technical Support* (CA9) given that collaborative infrastructure consisting of technology products may often necessitate maintenance and repair services.

The next section discusses the features presented by the various tools available in the technology arena. These features will be employed to respond to the aforementioned requirements.

B. System features overview

Nine requirement groups described in the previous section are met by various tools presented by commercial applications and academic researches. These tools are gathered in ten groups, labelled as *features* of CPD systems. Each feature is followed by its label.

Palacio et al. [14] state that technological infrastructure to meet the specified requirements should include features such as *communication service, mechanism to share and filter relevant information, mechanism to spot individual project progress, interaction mechanism for team members, status updates and tasks progress, search tool based on profile, status, and activity; synchronous and asynchronous communication*. PD oriented studies are also reviewed to support development process while technology planning. Sky and Buchal [26] categorize tools to support PD in six groups: *information gathering, drawing and design, analysis and evaluation, general documentation, planning and scheduling, synchronous workspace sharing*. Büyüközkan et al. [3] classify concurrent PD tools as *networking and management tools, modelling and analysis tools, predictive tools, and intelligent tools*.

Studies clearly emphasize the importance of communication tools. It is essential to assure coordination with ICT [1] and therefore communication tools are considered as primary features in a CPD system. Literature shows that synchronous and asynchronous communication tools are nearly always included in any collaborative system. *Synchronous communication tools* (EC1) assure real-time communication while spatially and temporally different communication is realized by *asynchronous communication tools* (EC2); which include e-mail, faxing, discussion boards, etc.

System integration mechanisms (EC2) are also widely studied in the literature. Some argue web-based interfaces to integrate various design models while others propose unification of modelling schemes [27]. *Project management tool* (EC4) is indispensable in a CPD project and it serves to control and coordinate the virtual team and their tasks [9]. *Product visualization* (EC4) is another feature of CPD systems. Collaborative visualization and collaborative design of the product allows teams to view, design, modify, mark-up, and measure the 3D virtual geometric model.

Document management tools (EC6) systems aim to store electronic documents and images, which enables engineering teams to create knowledge out of the information shared throughout the CPD project. *Content management tools* (EC7), serve to manage the workflow in collaborative environments.

Described as tools to keep track of history of a dataset [25], *Data Tracking & Analysis Tool* (EC8) enables the collaborating teams to make sense of the data they are handling. Data tracking is therefore important as it provides a detailed history of the data and the origin it generated from. *Archiving tools* (EC9) is also an important feature where large data is shared by distributed teams as storing, retrieving, and accessing the data are assured by archiving. It is important to be able to make use of the information created during the collaboration process. *Decision support tools* (EC10) become necessary at this stage, where a system is required to analyze all data and present an understandable report to assist decision makers' in their decision process.

Overall, ten system features are identified in response to the nine requirements of CPD projects.

V. IT PLANNING USING FUZZY QFD

Defining the requirements and the system features provides a better understanding of the current situation of CPD infrastructure. However, a planning methodology is required in order to map the aforementioned requirements into the features. HoQ diagram, the most recognized form of QFD, emerges as an appropriate planning tool. The translation of customer requirements into technical specifications becomes IT requirements for CPD mapped into CPD system features. Consequently, CAs are mapped into ECs in order to define how the system features respond to CPD requirements.

TABLE I. FUZZY SCALE FOR IMPORTANCE LEVELS

Scale for importance levels		
Linguistic variable	***Abbreviation***	***TFN***
Very low	VL	(0, 1, 2)
Low	L	(2, 3, 4)
Medium	M	(4, 5, 6)
High	H	(6, 7, 8)
Very high	VH	(8, 9, 10)
Scale for relationships		
Linguistic variable	***Abbreviation***	***TFN***
Strong	Θ	(7, 10, 10)
Moderate	O	(3, 5, 7)
Weak	▲	(0, 0, 3)
Scale for correlations		
Linguistic variable	***Abbreviation***	***TFN***
Strong positive	⊕	(3, 5, 7)
Positive	+	(0, 3, 5)
Negative	−	(-5, -3, 0)
Strong Negative	⊖	(-7, -5, -3)

Our expert; an e-Business specialist, Knowledge Management Group leader, and CRM coordinator; is consulted for his industrial insight on the importance of requirements, requirement-system feature relationships, and system features correlations. He is asked to evaluate domains 3, 4, and 5 according to the scales presented in Table I.

The HoQ evaluation is displayed in Fig. 2. The expert evaluation, contrary to expectations, covers all pairwise relationships in Domain 4 and all pairwise correlations in Domain 5.

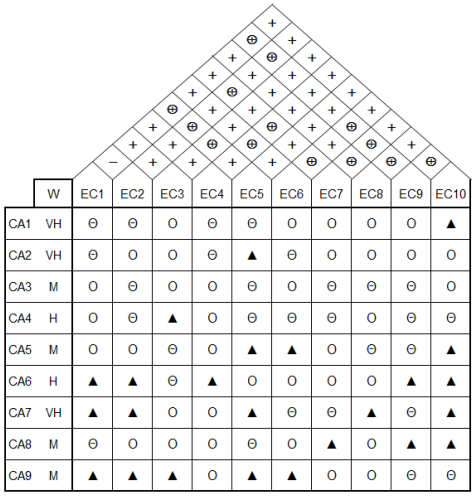


Figure 2 Expert judgments on weights, relationships and correlations

Fig. 2 lists the importance of the requirements. Then the mapping phase shows how these requirements are satisfied through the system features. Lastly, correlations between the system features are defined in order to observe their effect on each other.

These expert judgments are translated into TFNs according to the scales in Tables I. Priorities of system features are computed through equations (1) and (2). Final crisp priorities, displayed in Table II, are computed through equation (3). As a result, a priority vector is obtained for implementation of system features.

TABLE II. SYSTEM FEATURE PRIORITIES

	System Feature	Priorities (Normalized)
EC9	Archiving tools	0.122
EC7	Content management tools	0.119
EC5	Product visualization	0.106
EC6	Document management tools	0.105
EC8	Data tracking & analysis	0.098
EC4	Project management tool	0.098
EC10	Decision support tools	0.094
EC3	System integration mechanisms	0.090
EC2	Asynchronous communication tools	0.088
EC1	Synchronous communication tools	0.080

Final ranking of normalized priorities clearly suggest the importance of the archiving tools. This outcome can be interpreted as the importance of co-learning in CPD projects [1]. Archiving tools assure communication of the created information and sharing during collaboration efforts. Content management tools hold the second level of priority. This feature is also strongly connected with co-learning, whşch is a result of CPD. In a CPD project, product visualization tools rank as the third important system feature. This tool enables various engineering teams from various sites to conduct development process and therefore emerges as another high priority feature.

It is interesting to observe that communication tools (asynchronous and asynchronous) are the two system features with the least priorities. However, when combined, they emerge as the system feature with the most priority. This outcome can be linked to the fact that communication tools do not require high technology or high specification. Even the most basic communication tools can achieve the communication required in the CPD projects.

The outcome can be interpreted as an investment route for IT implementation at the beginning of a CPD project. This HoQ outcome aims to provide an understanding of implementation priorities from our expert's perspective.

CONCLUDING REMARKS

An essential part of CPD performance is IT; given that both the PD and the collaboration of various teams on different sites require a comprehensive technological infrastructure to support the communication as well as the integration of the firms. However, the rapid evolution of the IT complicates the forecasting and the planning of technological infrastructure.

The contribution of this study is two-fold. First, this paper proposes a set of technological requirements for CPD and a generic set of system features that includes tools and applications to respond to these requirements. On the other hand, a HoQ framework is employed to map the requirements to the system features. Importance of requirements, relationships between requirements and features, and correlations between system features included in the HoQ are evaluated by an industrial expert and Fuzzy QFD methodology is employed to interpret these evaluations.

Results show that system features associated with collaborative learning have the most priorities when planning technological infrastructure. However, it is apparent that all system features concur approximately to the same importance level. This can be interpreted as the need to cover all aspects of technological infrastructure within a CPD process. The outcome provides an implementation route for system features while considering IT infrastructure for CPD projects.

Further research includes extension of the current work through evaluations of different industrial experts in order to observe the differences in the priorities outcome according to industrial profile of the assessor. It is also anticipated to further develop the HoQ application in order to present a

 ISBN: 978-1-61208-176-2

comprehensive planning methodology, considering additional inputs.

ACKNOWLEDGMENT

Authors wish to thank TUBITAK for the financial support for this research, which is realized in the scope of TUBITAK Project No: 109M147. Authors also wish to thank Lemi Tuncer for his industrial insight during the development and evaluation processes.

REFERENCES

[1] J. Arsenyan and G. Büyüközkan, "Modelling Collaborative Product Development Using Axiomatic Design ," in CD proceedings of 15th International Conference on Concurrent Enterprising (ICE 2009) , Leiden, The Netherlands, 22-24 June 2009.

[2] T. Koc and Y. Mutu, "A Technology Planning Methodology Based on Axiomatic Design Approach," in PICMET 2006 Proceedings, Istanbul, Turkey, 2006, pp. 1450-1456.

[3] G. Büyüközkan, T. Dereli, and A. Baykasoğlu, "A survey on the methods and tools of concurrent new product development and agile manufacturing," Journal of Intelligent Manufacturing, vol. 15, pp. 731-751, 2004.

[4] A. L. Porter, A. T. Roper, T. W. Mason, F. A. Rossini, and J. Banks, Forecasting and Management of Technology. New York: Wiley, 1991.

[5] M.J. Martin, Managing innovation and entrepreneurship in technology-based firms. Canada: John Wiley & Sons, Inc., 1994.

[6] A. Rip and R. Kemp, "Technological Change," in Human Choice and Climate Change. Columbus, OH: Batele Press, 1998, pp. 327-399.

[7] M. W. Pretorius and G. de Wet, "A model for the assessment of new technology for the manufacturing enterprise," Technovation, vol. 20, no. 1, pp. 3-10, 2000.

[8] R. Kumar and P.S. Midha, "A QFD based methodology for evaluating a company's PDM requirements for collaborative product development," Industrial Management & Data Systems, vol. 101, no. 3, pp. 126-132, 2001.

[9] K. Rodriguez and A. Al-Ashaab, "Knowledge web-based system architecture for collaborative product development," Computers in Industry, vol. 56, pp. 125-140, 2005.

[10] G. Rueda and D.F. Kocaoglu, "Diffusion of emerging technologies: An innovative mixing approach," in PICMET 2008 Proceedings, Cape Town, South Africa, 2008, pp. 672-697.

[11] H. Shengbin, Y. Bo, and W. Weiwei, "Research on application of technology roadmap in technology selection decision," in Control and Decision Conference, 2008. CCDC 2008, Yantai, Shandong , 2008, pp. 2271 - 2275.

[12] D.-B. Luh, Y.-T. Ko, and C.-H. Ma, "A Dynamic Planning Approach for New Product Development," Concurrent Engineering, vol. 17, no. 1, pp. 43-59, 2009.

[13] Y.-T. Ko, "A dynamic planning method for new product development management," Journal of the Chinese Institute of Industrial Engineers, vol. 27, no. 2, pp. 103-120, 2010.

[14] R.R. Palacio, A. Vizcaino, A.L. Moran, and V.M. Gonzalez, "Tool to facilitate appropriate interaction in global software development," IET Software, vol. 5, no. 2, pp. 157-171, 2011.

[15] J.R. Hauser and D. Clausing, "The House of Quality," Harvard Business Review, vol. 66, no. 3, p. 63–73, 1988.

[16] C.G. Şen and H. Baraçlı, "Fuzzy quality function deployment based methodology for acquiring enterprise software selection requirements," Expert Systems with Applications, vol. 37, p. 3415–3426, 2010.

[17] S. Vinodh and S.K. Chintha, "Application of fuzzy QFD for enabling leanness in a manufacturing organisation," International Journal of Production Research, vol. 49, no. 6, pp. 1627-1644, 2011.

[18] S. Vinodh and S.K. Chintha, "Application of fuzzy QFD for enabling agility in a manufacturing organization: A case study," The TQM Journal, vol. 23 , no. 3, pp. 343 - 357, 2011.

[19] A. Lee and C.-Y. Lin, "An integrated fuzzy QFD framework for new product development," Flexible Services and Manufacturing Journal, vol. 23, pp. 26-47, 2011.

[20] H.-T. Liu, "Product design and selection using fuzzy QFD and fuzzy MCDM approaches," Applied Mathematical Modelling, vol. 35, p. 482–496, 2011.

[21] G.Z. Jia and M. Bai, "An approach for manufacturing strategy development based on fuzzy-QFD," Computers & Industrial Engineering, vol. 60, p. 445–454, 2011.

[22] W. D. Li and Z. M. Qiu, "State-of-the-art technologies and methodologies for collaborative product development systems," International Journal of Production Research, vol. 44, no. 13, pp. 2525-2559, 2006.

[23] J. Li and D. Su, "Support modules and system structure of web-enabled collaborative environment for design and manufacture," International Journal of Production Research, vol. 46, no. 9, pp. 2397-2412, 2008.

[24] W. Shen, Q. Hao, and W. Li, "Computer supported collaborative design: Retrospective and perspective," Computers in Industry, vol. 59, p. 855–862, 2008.

[25] E.S. Lee, D.W. McDonald, N. Anderson, and P. Tarczy-Hornoch, "Incorporating collaboratory concepts into informatics in support of translational interdisciplinary biomedical research," International Journal of Medical Informatics, vol. 78, no. 1, pp. 10-21, 2009.

[26] R.W.E. Sky and R.O. Buchal, "Modeling and Implementing Concurrent Engineering in a Virtual Collaborative Environment," Concurrent Engineering, vol. 7, no. 4, pp. 279-289, 1999.

[27] G. Buyukozkan and J. Arsenyan, "Collaborative Product Development: A Literature Overview," Production Planning & Control, vol. Accepted., 2010.

 ISBN: 978-1-61208-176-2

Indoor IEEE 802.11g Radio Coverage Study

Sandra Sendra, Laura Ferrando, Jaime Lloret, Alejandro Cánovas
Instituto de Investigación para la Gestión Integrada de zonas Costeras - Universidad Politécnica de Valencia, Spain
sansenco@posgrado.upv.es, laufermo@epsg.upv.es, jlloret@dcom.upv.es, alcasol@posgrado.upv.es

***Abstract*— Even though the wireless coverage inside buildings is widely studied, there are many deployments that are not optimal. An accurate design, not only allows more radio coverage inside the building, but could allow cost savings if we are able to reduce the number of access points that are used to implement the solution. In this paper, we will show a research study about the optimum location of access points inside the building of the Polytechnic University of Valencia (UPV) in order to provide better wireless Internet access to the students. The paper provides a comparative study for different Service Set Identifiers (SSIDs) (are currently available at the university). We will also compare them analytically. Finally we will obtain the mathematical expressions that allow us to model their behavior and we will see how the signals propagate following a very peculiar pattern.**

Keywords-radio coverage; indoor study, WLAN; IEEE 802.11g.

I. INTRODUCTION

Today, wireless networks are widely used in companies and universities as a support to the wired network. These allow the user easy access to all services provided by the company's network and usually allow Internet access.

When attempting to develop indoor wireless networks, many problems arise such as losses due to the walls, refractions over objects arranged randomly on the site or losses due to the use of different types of construction materials that cause different types of losses (brick, metal, glass, etc.) [1, 2].

Moreover to these issues, we should add the fact that the building interiors are almost never uniform making it very difficult to foresee in advance exactly what is the radio coverage of the wireless network [3].

It is necessary to have a correct and optimal design, in order to offer multiple additional services such as indoor positioning, object location, object tracking, etc. [4]. Moreover wireless coverage systems are being studied in other research fields such as Wireless Sensor Networks (WSN) [5].

In this paper, we analyze the behavior of the wireless signals from access points (APs) located in the Centre of resources for the research and learning (CRAI) of The Polytechnic High School of Gandia, of the Polytechnic University of Valencia (UPV). The obtained measurements will allow us to know the received signal strength evolution in function of the distance to the APs.

The rest of this paper is structured as follows. Section 2 shows some related works about radio coverage. Section 3 presents the scenario and the tools used to perform our measurements. Section 4 shows the result of the measurements in coverage maps. Section 5 makes a comparative study of the three analyzed signals on each floor. The analytical study and the equations which expressed this behavior are shown in Section 6. Finally, Section 7 shows the conclusion and future works.

II. RELATED WORKS

There have been many studies of wireless coverage, both empirically [6] and analytically [7]. Concretely, in [7], the authors performed an analytical study about the AP location and channel assignment. The treatment of these keys separately can lead to optimal designs. Authors propose an integrated model that addresses both aspects simultaneously in order to find a balance to optimize both objectives.

M. Kamenetskyt et al. [8] examine the methods for obtaining a position close to the optimal entry points and evaluate their performance in a typical center or campus environment. System performance is evaluated by an objective function, which aims to maximize the coverage area and signal quality. The optimization algorithms used are a subjective function on a discrete search space, which significantly reduces the complexity inherent to the problem. Numerical results show that random search algorithms, can lead to good solutions. However, the convergence of simulated travel speed depends largely on the development of simulation parameters and a good parameters selection.

Kaemarungsi and Krishnamurthy [4] studied the features of the IEEE 802.11-based WLANs and analyze the data in order to understand the underlying features of location fingerprints. They pointed out that the user's presence should be taken into account when collecting the location fingerprint for user related location-based service.

J. Lloret et al. [10, 11] showed studies about an empirical coverage radio model for indoor wireless LAN design. This model has been tested on a vast number of buildings of a great extension area with over 400 wireless APs in order to get the results successfully. The objective of the model is to facilitate the design of a wireless local area network WLAN using simple calculations, because the use of statistical methods takes too much time and it is difficult to implement in most situations. The proposed model is based on a derivation of the field equation of free propagation, and takes into account the structure of the building and its materials.

Sendra et al. [12] presented a comparison of the IEEE 802.11a/b/g/n variants in indoor environments in order to know which is the best technology. This comparison is made in terms of received signal strength indicator (RSSI), coverage area and measurements of interferences between channels.

 ISBN: 978-1-61208-176-2

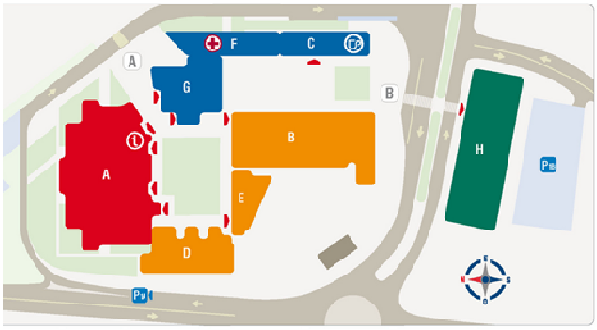


Figure 1. Map of Polytechnic high school of Gandia.

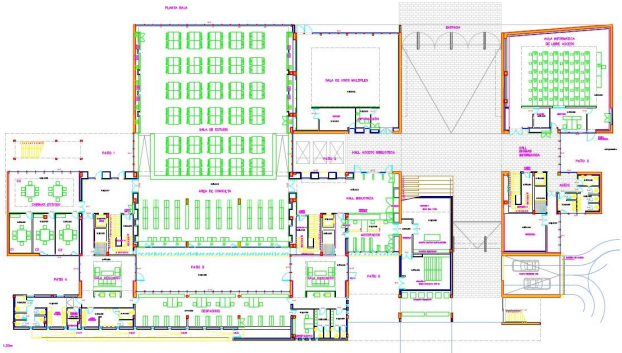


Figure 2. Ground floor of the CRAI bulilding.

Figure 3. First floor of the CRAI building

Figure 4. Second floor of the CRAI building

III. SCENARIO DESCRIPTION AND USED TOOLS

The CRAI was built in 2007. It belongs to the Polytechnic high school of Gandia. It has 3 floors, where different services for the students are offered. It contains the library, some computer labs and open access classrooms. Figure 1 shows the location of this space. It is the H building

Now, we are going to describe the scenario where the measurements have taken from the wireless networks and the hardware and software used to perform our research.

A. The building

The ground floor (see Fig. 2) contains the information desk, staff offices, the library and a large study room with a consultation area and several classrooms for group study. There is a multipurpose room where events and exhibitions are sometimes held.

On the first floor (see Fig. 3) we can find several computer labs, classrooms to perform Final Degree Projects, group study rooms and individual study rooms.

The second floor (see Fig. 4) has the magazine and journal library, the video library, some computer labs and some professor offices.

B. Description of UPV Wireless Network

Polytechnic high school of Gandia is a campus of the UPV and shares the 4 networks with the main campus: EDUROAM, UPVNET2G, UPVNET and UPV-INFO. Each one of these allows the university users to access Internet and the university resources.

- UPVNET: Wireless network with direct connection to all the resources of the UPV. It requires a wireless card with Wi-Fi Protected Access (WPA) and Wi-Fi Protected Access II (WPA2).
- UPVNET2G: Direct network connection to all resources of the UPV and the Internet. It requires a wireless card with WPA/WPA2.
- EDUROAM: This wireless network is widely deployed in universities and research centers in Europe. It provides internet access to all of their members. The users need the user name and password of their home institution. It requires a wireless card with WPA/WPA2. EDUROAM only provides Internet access.
- UPV-INFO: This wireless network is designed to provide information about how the wireless network should be configured. It uses private IP addressing and it does not allow Internet access. A second connection is needed to access Internet and UPV resources. This second connection is a Virtual Private Networking (VPN). It should only be used by very old computers that do not support WPA encryption.

In this paper, we will analyze three of these networks. They are UPVNET, UPVNET2G, EDUROAM.

C. Used software and hardware

In order to carry out this work, several measurements have been done along the three floors of the CRAI. We have used different network devices to perform the measurements:

- Linksys WUSB600N [13]: Is a USB wireless device that was used to gather the measurements. It is able to capture signals from the IEEE 802.11a/b/g/n variants. Its power transmission is 16 dBm for all variants and the receiver sensitivity is about -91dBm in both internal antennas. The transmission power consumption is less

 ISBN: 978-1-61208-176-2

than 480mA and it consumes 300mA in the reception mode.

- Laptop: It was used to take the coverage measurements. It has a dual core processor with 2 GHz per core and 2 Gbyte of RAM Memory. It has Windows Vista.
- Cisco Aironet 1130AG (AIR-AP1131AG-E-K9) [14]: This AP is used in all floors of the building. Its data rate can reach up to 54 Mbps. It can work at 2.4 GHz or 5 GHz, with a maximum distance between 100m to 122m indoors (as a function of the IEEE 802.11a or IEEE 802.11g variant). The maximum distance for outdoor environments, is between 198 m to 274 m. It can be powered by PoE (Power over Ethernet).

In order to capture the received signal at each point of the building, we used the following program:

- InSSIDer [15]: Is a free software tool that detects and controls the wireless networks and the signal strength in a graphical way. This program lists all detected wireless networks and provides their details as Service Set Identifier (SSID), Media Access Control address (MAC address), channel, RSSI, network type, security, speed and intensity of the signal and also allows the control of the quality of the signal.

IV. COVERAGE RESULTS

We have measured the walking area, where students and university staff can access. Bathrooms, exterior stairways, storage, etc are excluded. In order to perform this work, a grid of 4 meters x 4 meters has been drawn in each floor. This allows us to make measurements of the different networks in the same places. The laptop was located at a height of 100 cm above the ground.

A. Ground floor

This subsection shows the signal coverage measured on the ground floor. There are 5 APs that cover the entire plant. There are four places with the highest coverage level (the values are higher than -50 dBm). We highlight 2 rooms, Room A, the multipurpose room, and Room B, the computer room (see Figs. 5, 6, and 7). The AP located outside the wall of the computer room provides coverage levels below -70 dBm inside the classroom for all three cases.

Fig. 5 shows the coverage area and levels of UPVNET wireless network at the ground floor. Room A presents signal strength of -90 dBm due to the signal attenuation generated when crossing several walls.

Fig. 6 shows the coverage area for the UPVNET2G wireless network on the ground floor. Three places with higher signal strengths than -50 dBm can be seen. These places correspond to the location of APs. The multipurpose room has a very low coverage on the left side because the signal is greatly attenuated because it crosses several walls.

Fig. 7 shows the value of signal strength for EDUROAM wireless network on the ground floor. Again, there are three places with the high signal strength in excess -50 dBm, which correspond to the location of APs. In this case, more than half of the room B has signal strength levels below -70dBm.

B. First floor

This subsection shows the signal strength measured on the first floor. In this case, there are 4 APs that cover the entire plant. There are 4 places with the highest signal strength (higher than -50dBm).

Fig. 8 shows the signal strength for UPVNET wireless network on the first floor. The rooms at the left side have low radio coverage because the AP is not located in the correct place. The offices at the right side have also very poor signal strength because they are very close to the stairs and they suffer signal attenuation rather important.

Fig. 9 shows the UPVNET2G wireless network signal strength on the first floor. We can see that in the classroom on the left hand is not well covered because of the AP position. It is located on the right hand of the wall. The offices from the bottom right also have very poor coverage, because they are very close to the stairs, which generate significant signal attenuation.

Fig. 10 shows the EDUROAM signal strength on the first floor. In this case, we can see the same effect as in the other cases, but, moreover, there are tables in the study area (center of the picture), with a low signal strength (lower than -90 dBm).

C. Second floor.

This subsection shows the signal strength measured on the second floor. The floor is covered by 4 APs. They offer a good coverage across most of the surface.

Fig. 11 shows the signal strength for UPVNET wireless network on the second floor. In this floor, there is good radio coverage across the whole plant.

Fig. 12 shows the signal strength of UPVNET2G on the second floor. This signal is correctly broadcasted through the floor and its signal strength levels are sufficient to cover the working places.

Fig. 13 shows the signal strength of EDUROAM wireless network on the second floor. The signal on this floor is very good.

After having analyzed the radio coverage images, it is easy to see that the behavior of the signal strengths for each floor is quite similar, with small variations. The received signal strength is very low in bathrooms and toilets. This is because the amount of water pipes and copper tubes in the walls which produce the signal attenuation. We have also found low signal strength levels in the stairwells. The stairs usually have metal framework and a foundation which prevents the spreading of the signal.

V. COMPARATIVE STUDY

In this section, we compare the signals of the same plant. Fig. 14 shows the three signals from the ground floor. UPVNET2G provides better signal strength levels than UPVNET and EDUROAM. Signal strengths from the first floor are shown in Fig. 15. UPVNET2G is the network that presents the highest signal strengths. UPVNET and EDUROAM have similar behavior although there are some points where EDUROAM signal is better.

 ISBN: 978-1-61208-176-2

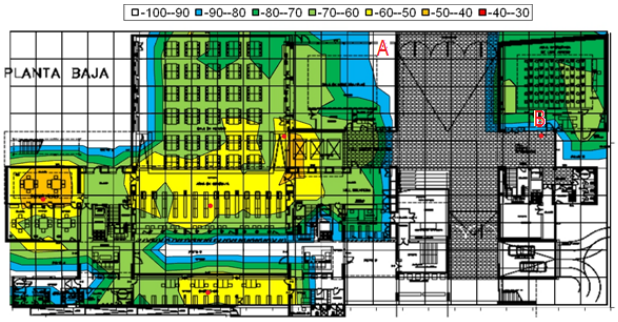


Figure 5. Radio coverage map of the ground floor for UPVNET

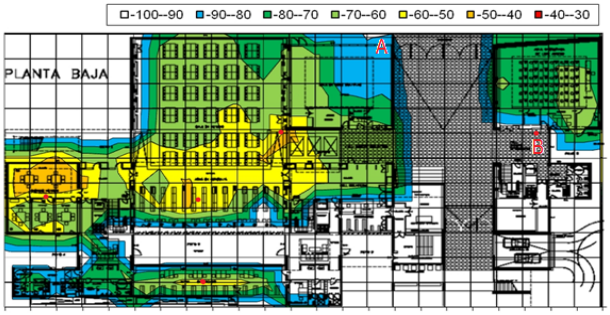


Figure 6. Radio coverage map of the ground floor for UPVNET2G

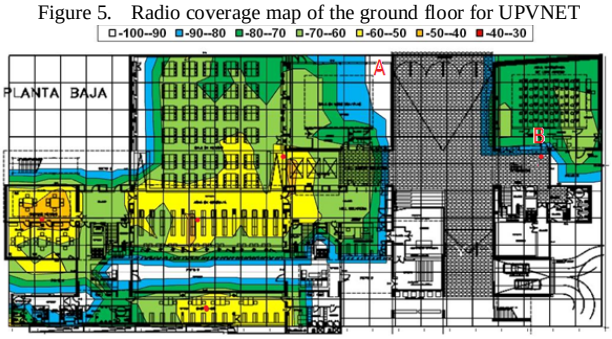


Figure 7. Radio coverage map of the ground floor for EDUROAM

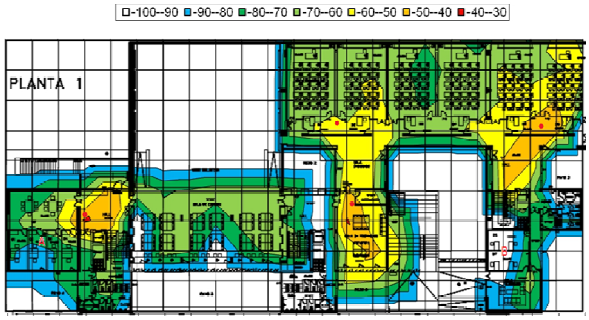


Figure 8. Radio coverage map of the first floor for UPVNET.

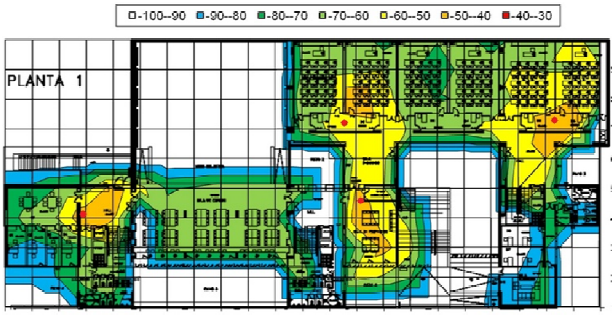


Figure 9. Radio coverage map of the first floor for UPVNET2G

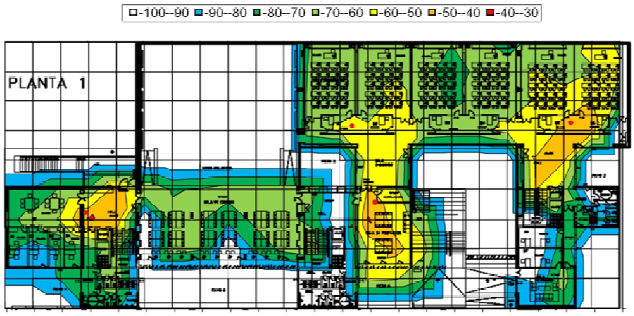


Figure 10. Radio coverage map of the first floor for EDUROAM

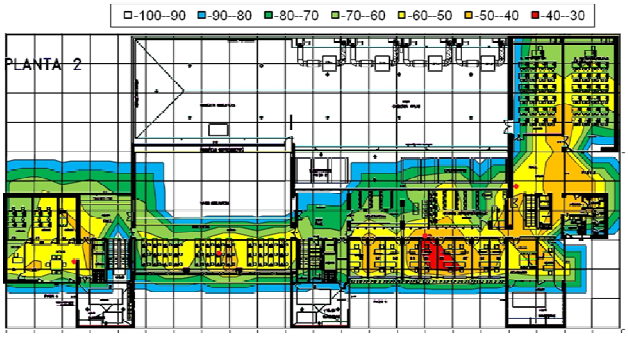


Figure 11. Radio coverage map of the second floor for UPVNET

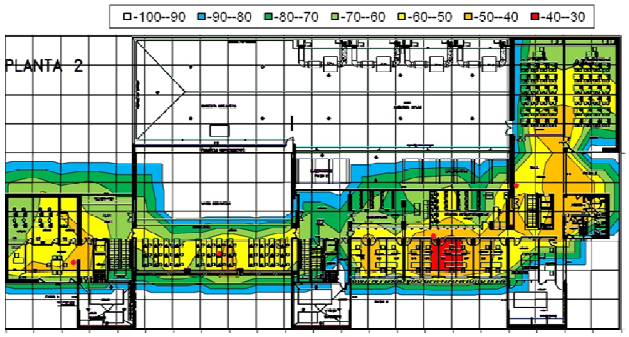


Figure 12. Radio coverage map of the second floor for UPVNET2G

Fig. 16 shows the behavior of signal strengths of the second floor. UPVNET2G and EDUROAM have identical behavior from 3 meters to around 10 meters, but from 0 to 3 meters and 10 meters to 12 meters, EDUROAM signal strength is better. The lowest signal strength is presented by UPVNET all the time. Keeping in mind all graphs, it is easy to conclude that the wireless network that provides the best signal strength level is UPVNET2G.

VI. ANALYTICAL STUDY

After analyzing the above figures, we can estimate the behavior of the wireless signals in indoor environments. Therefore, this section shows how the signal strength varies depending on the distance to the AP.

 ISBN: 978-1-61208-176-2

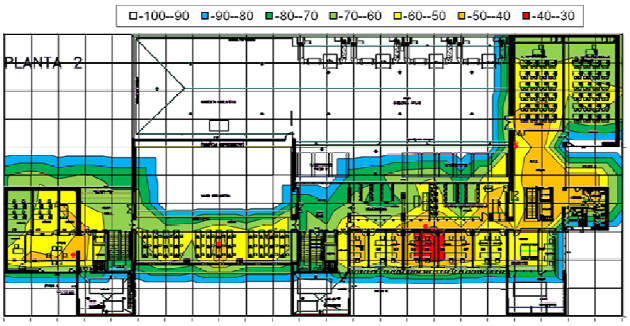


Figure 13. Radio coverage map of the second floor for EDUROAM

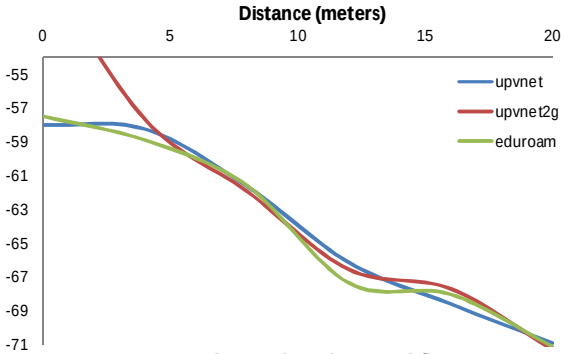


Figure 14. Signal strength on the ground floor.

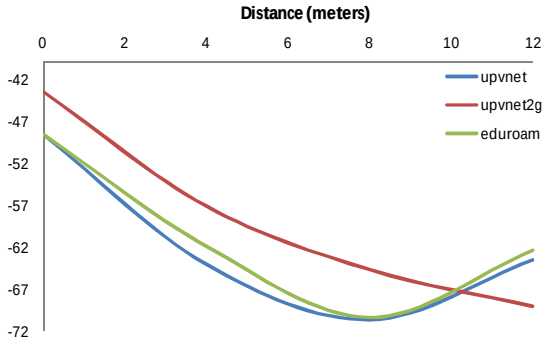


Figure 15. Signal strength on the first floor.

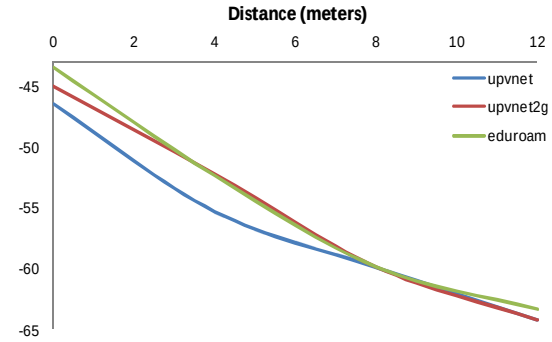


Figure 16. Signal strength on the second floor.

The analytical study is performed for three networks (UPVNET, UPVNET2G and EDUROAM). In order to draw each one of these graphs, we have estimated the average value of the three signals provided by each wireless network.

Fig. 17 shows the average value of the signal strength depending on the distance to the AP on the ground floor. Expression 1 shows the equation for the trend line (black line in Fig. 17) of our measurements. As we can see, it is a polynomial expression of fifth grade, with a correlation coefficient of $R^2=1$. However, we can appreciate a slight difference between them in positions close to 3-4 meters, and further away than 17 meters of the APs.

$$Y=-0.0001x^5+0.0066x^4-0.1078x^3+0.6889x^2-2.3012x-54.75 \quad (1)$$

where *Y* represents the average value of the received signal strength in dBm and X is the distance in meters to the AP.

Fig. 18 shows the average signal strength provided by the APs located on the first floor, as a function of the distance to the APs. In positions further than 8 meters of the APs, both graphs vary very few between them, although the rest of the graph is identical. Equation 2 shows the expression for the trend line (black line in Fig. 18) of our measurements. The behavior of wireless signals based on the distance is described by a cubic polynomial with a correlation coefficient of $R^2=1$.

$$Y=-0.0117x^3+0.0665x^2-3.9909x-46.833 \quad (2)$$

where Y is the signal level in dBm and X is the distance in meters to the AP.

In Fig. 19, provides the behavior of the signal level on the second floor. Equation 3 shows the trend line (black line in Fig. 19) of our measurements. In this case equation 3 is a 3rd degree polynomial with a correlation coefficient of $R^2=1$. It shows that both graphs have a nearly perfect match, as its correlation coefficient shows.

$$Y = 0.0021x^3 + 0.0292x^2 - 2.2229x - 45 \quad (3)$$

where Y is the average signal value in dBm and X is the distance in meters to the AP.

VII. CONCLUSION

When the deployment of a wireless network on the inside a building is needed to offer complete coverage for all the users, we should pay particular attention to the correct placement of the AP. We have analyzed the behavior of the signal strengths of the APs located in the CRAI. The measurements provided have enabled us to represent the signal evolution depending on the distance to the AP.

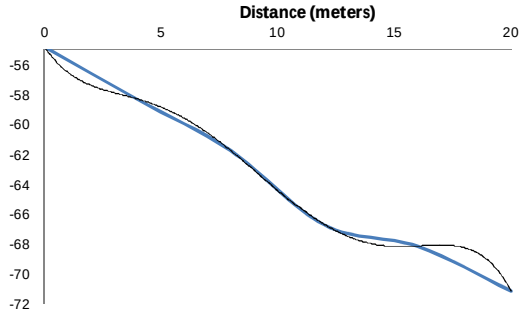


Figure 17. Average signal strength of the ground floor

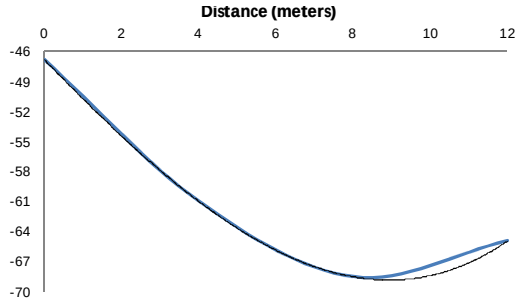


Figure 18. Average signal strength of the first floor.

With these measurements of the coverage maps of each floor, have been drawn the pictures processed for this analysis have allowed us to determine the places where the signal is not good (less than -70 dBm), and so relocate APs and add more, if it was needed. With all of these, we have performed an analytical and comparative study, with the three networks. After this study, we can see that, the EDUROAM and UPVNET have very similar coverage, while UPVNET2G is a little bit better. This phenomenon is a strange because the same APs give the coverage for the 3 wireless networks. Moreover, we have detected that the worst designed floor was the ground floor (in terms of APs distribution). We propose the relocation of some of the existent APs and to add new APs to cover the shadow areas where the signal strength is below than levels appropriate for Internet access.

We have characterized mathematically, the behavior of the signal strength in each floor. In all cases, the behavior can be expressed by a polynomial expression of degree equal to or greater than 3, depending on each floor. The APs of the CRAI building give acceptable radio coverage levels up to 16 meters from the AP's position.

On the other hand, we have estimated the average value for all floors, depending on the distance to the APs and we see that it could be modeled as a fifth degree polynomial with a correlation coefficient of 1. It is shown in equation 4. The signal strength Y is given in dBm and X the distance to the AP, in meters.

$$Y = 4 \cdot 10^{-5} x^5 - 0.0024 x^4 + 0.0455 x^3 - 0.2065 x^2 - 1.966 x - 50.792 \quad (4)$$

Finally, we have observed a trend in the signal behavior where it reduces its strength in a staggered manner (as we see clearly in Fig.14 and 17). In all other plants, this behavior is less visible. However, due to new tests that we are carrying out in other buildings, it seems that this pattern is repeated.

We are now working on the design to place more than one AP in the same location in order to provide higher bandwidths for the students. Moreover, we are proposing the use of APs in standby mode to provide fault tolerance. Finally, APs should be updated to IEEE 802.11n standard in order to achieve higher speeds and greater distances.

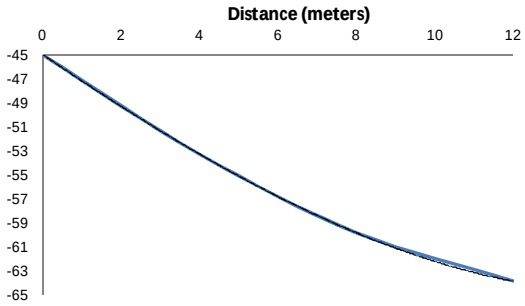


Figure 19. Average signal strength of the second floor

REFERENCES

[1] J. Lloret, J. J. López, and G. Ramos, "Wireless LAN Deployment in Large Extension Areas: The Case of a University Campus", In proceedings of Communication Systems and Networks 2003, Benalmádena, Málaga (España), September 8-10, 2003.

[2] N. Pérez, C. Pabón, J. R. Uzcátegui, and E. Malaver,"Nuevo modelo de propagación para redes WLAN operando en 2.4 Ghz, en ambientes interiores". TÉLÉMATIQUE 2010, Vol.9, Issue: 3, pp.1-22.

[3] B. S. Dinesh, "Indoor Propagation Modeling at 2.4 Ghz for IEEE 802.11 Networks". Thesis Prepared for the Degree of master of science university of north texas. December 2005

[4] K. Kaemarungsi and P. Krishnamurthy, "Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting". In proceedings of The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services 2004 (MOBIQUITOUS 2004) . Boston, Massachusetts, USA , August 22-26, 2004.

[5] R. Mulligan and H. M. Ammari, "Coverage in Wireless Sensor Networks: A Survey", Journal of Network Protocols and Algorithms. Vol. 2, No. 2, 2010.

[6] A. R. Sandeep, Y. Shreyas, S. Shivam, A. Rajat, and G.Sadashivappa, "Wireless Network Visualization and Indoor Empirical propagation Model for a Campus Wi-Fi Network ", Journal of World Academy of Science, Engineering and Technology, Vol 42, Pp. 730-734, 2008.

[7] A. Eisenbl, H-F. Geerdes, and I. Siomina, " Integrated Access Point Placement and Channel Assignment for Wireless LANs in an Indoor Office Environment",In proceedings of IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks(WoWMoM 2007). Helsinki, Finland, June 18-21, 2007

[8] M. Kamenetskyt and M. Unbehau, "Coverage Planning for Outdoor Wireless LAN Systems", International Zurich Seminar on Broadband Communications 2002, Zurich (Switzerland), February, 19-21, 2002.

[9] IEEE Std 802.11 (2007) IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Pp.1-1184. New York, USA.

[10] J. Lloret, J. J. López, C. Turró, and S. Flores, "A Fast Design Model for Indoor Radio Coverage in the 2.4 GHz Wireless LAN", in proceedings of 1st Int. Symposium on Wireless Communication Systems 2004 (ISWCS'04), Port Louis (Mauricio Island), September 20-22, 2004.

[11] J. Lloret and J. J. López, "Despliegue de Redes WLAN de Gran Extensión, el Caso de la Universidad Politécnica de Valencia", XVIII Simposium Nacional de la Unión Científica Internacional de Radio, ACoruña (Spain), September 10-12, 2003.

[12] S. Sendra, P. Fernandez, C. Turro, and J. Lloret, " IEEE 802.11a/b/g/n Indoor Coverage and Performance Comparison", In procedings of the 6th International Conference on Wireless and Mobile Communications (ICWMC 2010), September 20-25, 2010, Valencia, Spain.

[13] Data sheet WUSB600N. available. in: http://www.linksysbycisco.com/EU/es/products/WUSB600N <retrieved: Nov.,2011>

[14] Specifications of cisco Aironet 1130 AG, Access Point. Avaliable in: http://www.cisco.com/en/US/prod/collateral/wireless/ps5678/ps6087/product_data_sheet0900aecd801b9058.html <retrieved: Nov., 2011>

[15] Web page of inSSIDer available in: http://www.metageek.net/products/inssider <retrieved: Nov., 2011>

 ISBN: 978-1-61208-176-2

Security Issues of WiMAX Networks with High Altitude Platforms

Ilija Basicevic, Miroslav Popovic
Faculty of Technical Sciences
University of Novi Sad
Novi Sad, Serbia
ilibas@uns.ac.rs, miroslav.popovic@rt-rk.com

***Abstract*—In this paper, we discuss the possibility of securing High Altitude Platform Networks (HAP) networks with intrusion detection systems (IDS). We assume that it is a 802.16 network, in point-to-multi point mode. An analysis of possible threats and attack sources is given. Based on that analysis and specific properties of HAP networks an IDS concept is proposed. The main idea of the concept is that a network-based IDS system is collocated with the base station (BS) software. The BS is on board the HAP. The extensions to the concept, in order to provide for prevention feature, are outlined. In that case, the correlation module is a publish/subscribe server for dissemination of events that are the results of alert correlation. Its subscribers are policy enforcement points in the HAP network.**

Keywords-high altitude platforms; IEEE 802.16; intrusion detection system; network security

I. INTRODUCTION

In the recent years, there has been a strong incentive in research of High Altitude Platform (HAP) networks. While in the first place they were envisioned as a means for rapid provisioning of connectivity in the case of disasters (because of the short time needed to launch a HAP vehicle and establish connectivity), soon other scenarios have been proposed as well. For example, in rural areas, with scarce or not existing ground infrastructure, HAPs can be used to provide broadband connectivity, see Fig 1. Another possible application is in mobile sites (e.g., trains). Applications in military communications are also considered. Researchers envisioned high-rate communications (up to 120 Mb/s) delivered directly to a user in line of site of a HAP within a coverage area up to 60 km wide [6]. There are three expected scenarios regarding the position of HAP in the end-to-end path [6]:

- Isolated from any core networks, providing connectivity for private networks
- Between core networks as point-to-point trunk connections
- In the access network, providing users with access to core networks

The significance of first-responder communications (e.g., Enhanced 911 service in US) during catastrophic events is utmost. Such services can be provided by dispatching HAP vehicle with telecommunications equipment in the affected area.

HAP is defined as a solar-powered unmanned airship or airplane, capable of long endurance on-station (several months or more) [7]. The HAP payload can be a complete base station. Besides up- and down-links to the user terminals, and backhaul links, links to satellites can be established as well. In some scenarios, where networks of HAPs are applied, there are also inter-HAP links.

The coverage region is determined by line-of-sight propagation and the minimum elevation angle at the ground terminal. The advantages of HAP communications are [7]:

1. Large area coverage (compared with terrestrial systems)
2. Flexibility to respond to traffic demands - flexible and responsive frequency reuse patterns and cell sizes, unconstrained by the physical location of base-stations.
3. Low cost - cheaper to launch than a geostationary satellite or a constellation of Low Earth Orbit (LEO) satellites, cheaper to deploy than a terrestrial network.
4. Incremental deployment - service may be provided initially with a single HAP and expanded gradually - in contrast to LEO satellites.
5. Rapid deployment - it is possible to design, implement and a deploy HAP service relatively quickly, especially when compared to satellites.
6. Platform and payload upgrading - can be relatively easily and safely brought down for payload upgrading.
7. Environmentally friendly

The backhaul link is realized using cellular scheme too, because a single link can not provide full backhaul capacity. Thus there are going to be a number of distributed backhaul ground stations, though this number can be fewer than the number of user cells served because of the higher order modulation schemes that would be used in backhaul links, which would provide greater capacity [7].

HAP-based services have been allocated frequencies by the ITU at 47/48 GHZ, also at 28 GHz in ITU Region 3 - Asia.

Most of the scenarios predict use of HAPs for 802.16 networks, although Universal Mobile Telecommunications System (UMTS) is present in application scenarios as well, albeit in much smaller extent. We assume that 802.16 network is in point-to-multi point (PMP) mode. With regard to physical characteristics of the network, HAP is usually positioned at an altitude of approximately 17-22 kilometers. It covers up to 256 cells.

 ISBN: 978-1-61208-176-2

Use of HAP platforms for different applications has been studied in the scope of several projects (HAPCOS – EU COST action 297, HELINET and CAPANINA EU Framework Programme projects), but to the best of our knowledge this is the first analysis of the possibilities for protection of HAP WiMAX networks with IDS systems.

Section 2 briefly presents security mechanisms that are used in 802.16 networks. Section 3 describes architecture of a network based intrusion detection system for 802.16 networks. It includes analysis of possible threats and possible improvements in order to realize prevention feature (besides detection). Section 4 contains concluding remarks.

II. SECURITY MECHANISMS IN 802.16 NETWORK TECHNOLOGY

Compared to IEEE 802.11, a serious effort has been undertaken in designing the security mechanisms in IEEE 802.16. The following description is based on [1].

IEEE 802.16 protocol stack contains Media Access Control (MAC) layer, which is divided into three sublayers (convergence sublayer, common part sublayer and privacy sublayer). Service specific convergence sublayer has two types, one that interfaces ATM as upper layer, and the other for TCP/IP. Common part sublayer is the core part of IEEE 802.16 MAC. It manages connections and bandwidth, among other functions. There are three types of connections: Primary, Basic and Secondary. Primary are used for authentication. Basic are used for time critical MAC control messages. Secondary are used for standards based management messages (e.g., SNMP [5]). MAC is connection oriented, and all data communications are in the context of connection. Connections are added, modified and deleted dynamically. Privacy sublayer is responsible for security functions:

- Encryption,
- Decryption,
- Authentication,
- Secure key exchange.

This sublayer contains two protocols: Encapsulation and the Privacy and Key Management Protocol (PKM).

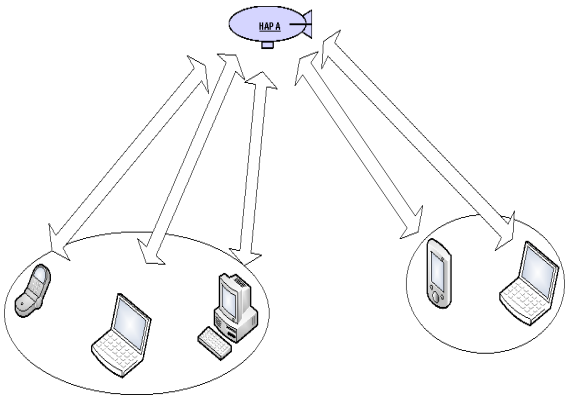


Figure 1. HAP network

Security protocols use Security Associations (SA). There are three types of SAs: primary, static and dynamic. Primary are established during initialization process. Static and dynamic can be shared between different subscriber stations (SS). Shared information may include traffic encryption key (TEK) and initialization vector (IV).

BS is responsible for maintaining keying information for all SAs. SA keying material has limited lifetime. When BS provides SS with a keying material, it includes information on remaining lifetime. Keying system is a two-tier one. The first tier is: using public keys, BS sends authorization key (AK) to SS. The second tier is that by using AK, the TEK exchange is protected.

PKM protocol is used for synchronization of keying information between BS and SS. PKM has two finite state machines: Authorization and TEK exchange. PKM authorization is realized as an exchange of three messages. In this exchange, SS provides BS with its certificates (certificate of the manufacturer and of the station itself) - BS authenticates SS, and BS provides SS with AK and with identification of SAs it is authorized to access. Key encryption key (KEK) and message authentication keys are derived from AK. Security components use X.509 Version 3 certificates. PKM TEK exchange is an exchange of two or three messages, in which BS sends TEK parameters for each requested SA. Three messages are exchanged. PKM is a client/server protocol where SS is a client. PKM uses RSA with SHA-1. IEEE 802.16 encryption uses DES CBC over payload. Generic MAC Header (GMH) and CRC fields are not encrypted.

New subscriber station enters the network in five steps:

1. SS scans for a BS downlink signal and uses it to establish channel parameters
2. Primary management connection established
3. SS authorized using PKM
4. SS sends a register request and BS responds with second management connection Id
5. Transport connections are created

The first phase is security capabilities negotiation during which SS informs BS which cryptographic suites it supports and BS tells SS which of those to use in the subsequent communication (this information is contained in the descriptor of the primary SA).

BS generates AKs and TEKs using random or pseudo-random generators. IVs are generated in such a manner as to be unpredictable.

III. NIDS FOR IEEE 802.16 - HAP CASE

Network-based IDS are an important class of contemporary IDS systems. In this class of systems, IDS processes the stream of packets that are transmitted over the network. In the typical wired environment IDS is usually placed in the network perimeter. In the HAP network, which is a 802.16 network in the point-to-multi point (PMP) mode, a logical place for NIDS is the base station. We propose architecture with a NIDS sensor in each cell, a NIDS sensor that monitors link to the gateway and a coordination and correlation module which has the following functions:

 ISBN: 978-1-61208-176-2

- It correlates alerts from NIDS sensors in 802.16 cells
- It sends the results of the correlation phase to the network operations center
- It downloads signature updates from public repository or vendor web site
- It uploads signature updates to NIDS sensors

Snort is a popular IDS system, significantly present both in everyday use, and in research. It is an IDS system for wired networks, but there are important common concepts that IDS systems for wired and wireless networks share. The Snort 3.0 architecture [8] promotes separation of the Snort Security Platform (SnortSP) from the Engines module which contains analytics modules. We propose that Snort 3.0 architecture can be used as IDS architecture for HAP systems. In the HAP case, the Dispatcher module connects not only to the local Data Source, but also receives alerts from the coordination module.

SNMP [5] is a de-facto standard for network management in Internet environment, thus we propose that in HAP IDS system the same protocol would be used. In case of HAP networks, it is usually assumed that the network operations center is placed on the ground, and that a wireless link is used for network management of the HAP network, for software updates and maintenance and also in some cases for the maneuvering of the HAP vehicle. The network operations center contains the Security Officer console, which allows for visual inspection of the state of HAP IDS, the list of most recent alerts and similar features.

A. Threats to the HAP network

As a primary attack venue, we see subscriber stations. We divide the attacks into three classes:

- Attacks at the physical layer,
- Attacks at the MAC layer,
- Attacks at higher layers.

Subscriber stations in the HAP cell can mount physical layer attacks. In the literature are often mentioned jamming and packet scrambling, belonging to this class. The jamming attack is mounted using information from UL-MAP message received from BS, and if it is a targeted attack, attacker has to map the CID from UL-MAP message to the station address. This attack can be realized with short transmissions and low radiated power, which protects the attacker [11].

At MAC layer, subscriber stations are capable of mounting Denial of Service (DOS) attacks. DOS attacks at MAC layer are realized as flooding of signalization requests (authentication, capabilities negotiation, key management frames, etc.). The primary means of those attacks are resource intensive cryptographic operations.

Some of messages in IEEE 802.16 are not authenticated (Traffic Indication Message, Neighbor Advertisement Message, Fast Power Control, Multicast Assignment Request, Downlink Burst Profile Change Request, Power Control Mode Change Request) [12] which leaves space for attacks.

At higher layers, a distributed variant of DOS (DDOS) attacks is possible. DOS attacks at application layer that disturb the normal application operation instead of depleting network resources as in classical DOS, are becoming more and more serious threat recently. At application layer, also are possible targeted attacks at users. Typical examples are different types of malware hidden in email attachments. Protective measures include application level filters at network servers. Those are outside the scope of this paper.

One often cited type of attack, which is possible in 802.16, although it is more difficult to realize than in 802.11 is the rogue base station attack. This type of attack belongs to the class of man-in-the middle attacks. In the attack, the rogue base station impersonates a legitimate one. A short description of the attack is given in [4]. Other attacks in this class are more probable in a mesh network, rather than in a network in PMP mode. The proposed IDS system at this moment does not include the detection feature for this type of attacks.

Besides subscriber stations, the source of attacks in higher layers of protocol stack can be in external networks - mounted over the link to the gateway. Those attacks are targeted at stations in the HAP network. As this is the last hop in the communication path between attacker and its target, DOS attacks are already amplified and easily detectable, but the possibility for reaction is limited.

The last is that although SNMPv3 includes authentication, it has to be noted that the link to operations center presents another attack venue. The privileges that are given to management personnel are wide: software updates, installation of software modules, restart, power on/off, maneuvering in case of HAP airplanes, etc. Since the operations that are realized over the management interface are of great security impact, the damage that an attacker who successfully impersonates the network operations center could make is critical.

B. Remarks on the construction of HAP IDS

The first phase of detection in a NIDS is the packet "sniffing". While relatively simple for realization in a wired network, in wireless networks the NIDS system has to scan traffic at a set of frequencies. Each of the frequencies is scanned in specific intervals of time. Typically not the same time interval is devoted to scanning of all frequencies in the set, and there is usually a heuristic algorithm (sometimes based on fuzzy logic) applied to determine how long to scan each of the frequencies. The integration of IDS sensor software with the BS protocol stack software would provide for the simple method of monitoring of the communication between subscriber stations and the base station. The concept of integration can be similar to the use of filter hooks [9] and filtering platform callout drivers [10] in Microsoft Windows OS in packet filtering applications for wired networks. Fig 2 presents the structure of HAP IDS/IPS at one BS, including the information flows.

Average traffic load on HAP BS can be estimated in the following way. A traffic stream from one mobile user to the HAP BS can be modeled as 4IPP [13] (traffic model for IEEE 802.16.3) . Number of terrestrial users is 240-256 per cell in published simulations [14]. Thus, the total traffic on

 ISBN: 978-1-61208-176-2

HAP BS coming from terrestrial users in one cell in average case can be modeled as 256 4IPP streams. The HAP IDS should be able to inspect such a stream, without losing packets. The 4IPP average rate is 3 pkts/unit-of-time. The bandwidth of SS-HAP link is 1 Mbps in simulations [15]. The packet size is 1500 Bytes. The packet rate is 1000000/(1500*8), which is 83 packets per sec. The parameters of the basic 4IPP model (see Table 1, the 4IPP Average Rate is calculated as a sum of IPP stream rates and equals to 3 pkts/unit-of-time) should be scaled by 83/3=27.8 unit-of-time per sec. The resulting parameters of IPP streams for model of communication in HAP network are given in Table 2.

TABLE 1. PARAMETERS OF IPP STREAMS IN BASIC 4IPP TRAFFIC MODEL

Source #i	λi IPP in ON state (pkts/unit-of-time)	Averaged over both ON and OFF states (pkts/unit-of-time)
IPP#1	2.679	1.1480
IPP#2	1.698	.7278
IPP#3	1.388	.5949
IPP#4	1.234	.5289

TABLE 2. PARAMETERS OF IPP STREAMS IN HAP WIMAX TRAFFIC MODEL

Source #	Averaged over both ON and OFF states (pkts/sec) – for 1 SS	Averaged over both ON and OFF states (pkts/sec) – for 256 SS
IPP#1	31.91	8168.96
IPP#2	20.23	5178.88
IPP#3	16.54	4234.24
IPP#4	14.7	3763.2

Since it aggregates events from several cells, in some cases, the correlation module can detect low volume DDOS attacks (at higher layers of protocol stack), that would otherwise (without the aggregation and correlation of alerts coming from different cells) pass unnoticed.

In case of multi-HAP network, operation of HAP IDS systems belonging to specific HAP networks can be coordinated in centralized manner (from the network operations center on the ground), or those can cooperate in a distributed manner. The realization of such a cooperative system is outside the scope of this paper.

Attacker location is an important feature in wireless security. Application of techniques such as triangulation for that purpose is outside the scope of this paper.

C. *Possible improvements*

Besides the aforementioned cooperation in case of multi-HAP networks, there are two directions in which the proposed concept can be improved and/or extended.

The first one is that the 802.16 network can be used in mesh mode. There are already some proposals for IDS systems for wireless mesh networks: OpenLIDS [2], WATCHERS, TIARA, CONFIDANT, MobIDS, RESANE, SCAN [3]. The change from PMP to mesh mode would require a substantial rework of the concept.

The other direction is that having provided the detection functionality, the next step is the reaction feature (intrusion prevention). Such architecture is based on the use of Policy Enforcement Point engine (PEP) at the base station. It is often implemented as a firewall. In that case the functionality of the correlation module would be extended with the following function: dispatching of new alerts that are the results of correlation phase back to NIDS sensors. In order to achieve efficient use of communication and processing resources, the correlation module should be able to filter the alerts that it sends to sensors. There are strict limitations with respect to the weight of the load that can be placed in the HAP that imply the efficient use of processing resources. For that reason we propose that the coordination and correlation module is a lightweight topic-based publish/subscribe system. There is a publish/subscribe association between this module and PEP engines. In this association, the correlation module is the publisher and PEP engines are subscribers. We remark that in this design the PEP engines are collocated with sensors.

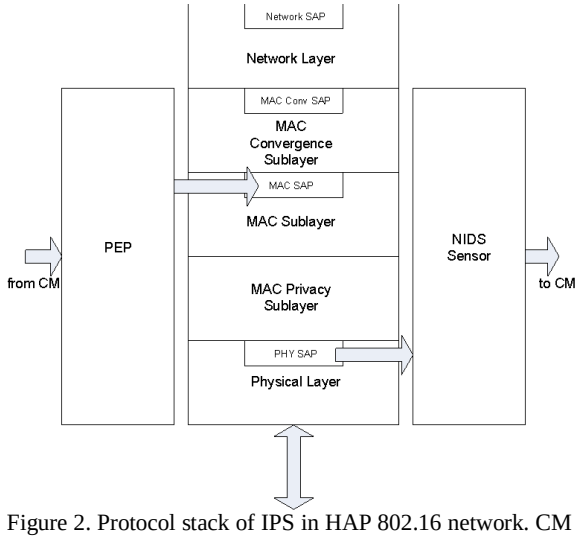


Figure 2. Protocol stack of IPS in HAP 802.16 network. CM is correlation module.

The proposed system is a good platform for realization of the reaction feature, because the communication stream from the correlation module to the PEP engines, which carries

results of correlation phase, in the average case contains enough information for decisions on reaction.

By implementing IDS/IPS as described, we allow for swift reaction in case of handovers of malicious subscriber stations. Once recognized as malicious, such a station can be disconnected and prevented from moving to the neighboring cell. This is especially of interest in overlapping areas, where mobile station can choose one of up to three cells (in average case) that it will use for communication.

IV. CONCLUSION AND FUTURE WORKS

This paper presents an approach to securing of HAP networks by using intrusion detection systems. The protected network is a 802.16 network in point-to-multi point mode. The proposed system is a distributed network-based IDS system with a NIDS sensor in each cell. IDS system is collocated with the base station software. The BS is on board the HAP. IDS sensors monitor communications links between base station and subscriber stations. The backhaul link to ground station is monitored as well. The IDS is collocated with the base station device, on board the HAP vehicle.

The system allows for detection of distributed DOS attacks in the HAP network. The paper describes the required modifications to the system in order to include reaction feature (intrusion prevention) in a straightforward manner. The correlation module in the extended system is the publish/subscribe server that publishes results of the correlation phase to the policy enforcement point engines in HAP cells.

The system could be further developed to include support for cooperation of IDS/IPS systems in multi-HAP networks.

ACKNOWLEDGMENT

This paper is a continuation of the research conducted in the scope of HAPCOS project (COST action 297). This work was partially supported by the Ministry of Education and Science of the Republic of Serbia under the project No. 32031 and 44009, year 2011.

REFERENCES

[1] IEEE Std 802.16™-2009, IEEE Standard for Local and metropolitan area networks, Part 16: Air Interface for Broadband Wireless Access Systems, IEEE-SA Standards Board, The Institute of Electrical and Electronics Engineers, Inc.

[2] F. Hugelshofer, P. Smith, D. Hutchison, and N.J.P. Race, OpenLIDS: A Lightweight Intrusion Detection System for Wireless Mesh Networks, Mobicom 09, Beijing, China, 2009, pp. 309-320

[3] T.M. Chen, G.S. Kuo, Z.P. Li, and G.M. Zhu, Intrusion Detection in Wireless Mesh Networks, chapter in Security in Wireless Mesh Networks, Auerbach Publications, 2008

[4] M. Barbeau, J. Hall, and E. Kranakis, Detecting Impersonation Attacks in Future Wireless and Mobile Networks, Workshop on Secure Mobile Ad-hoc Networks and Sensors, MADNES 2005

[5] U. Blumenthal and B. Wijnen, User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3), RFC 3414, 2002, The Internet Society

[6] D. Grace, M. Mohorcic, M.H.Capstick, M. Bobbio Pallavicini, and M. Fitch, "Integrating Users into the Wider Broadband Network via High Altitude Platforms", IEEE Wireless Communications, Vol. 12, No. 5, pp. 98-105, October 2005

[7] T.C. Tozer and D. Grace, "High Altitude Platforms for Wireless Communications", IEE Electronics and Communications Engineering Jnl, Vol. 13, No. 3, June 2001, pp. 127-137

[8] Snort 3.0 Architecture Series Part 1: Overview, http://securitysauce.blogspot.com/2007/11/snort-30-architecture-series-part-1.html, retrieved: November, 2011.

[9] Filter-Hook Drivers, http://msdn.microsoft.com/en-us/library/windows/hardware/ff546489%28v=vs.85%29.aspx, retrieved: November, 2011.

[10] Introduction to Windows Filtering Platform Callout Drivers, http://msdn.microsoft.com/en-us/library/ff556954%28VS.85%29.aspx, retrieved: November, 2011.

[11] Security of IEEE 802.16, Arkoudi-Vafea Aikaterini, Master Thesis, Department of computer and Systems Science, Royal Institute of Technology, Stockholm, Sweden, 2006

[12] A. Deininger, S. Kiyomoto, J. Kurihara, and T. Tanaka, “Security Vulnerabilities and Solutions in Mobile WiMAX”, IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.11, November 2007.

[13] C. R. Baugh, 4IPP Traffic Model for IEEE 802.16.3, IEEE 802.16.3c-00/51,

[14] Floriano De Rango, Mauro Tropea, and Salvatore Marano, Integrated Services on High Altitude Platform: Receiver Driven Smart Selection of HAP-Geo Satellite Wireless Access Segment and Performance Evaluation, International Journal of Wireless Information Networks, Vol. 13, No. 1, January 2006, pp. 77-94, DOI: 10.1007/s10776-005-0020-z

[15] C. E. Palazzi, C. Roseti, M. Luglio, M. Gerla, M. Y. Sanadidi, and J. Stepanek, Enhancing Transport Layer Capability in HAPS-Satellite Integrated Architecture, Wireless Personal Communications, Vol 32 Issue 3-4, February 2005

 ISBN: 978-1-61208-176-2

Alteration Method of Schedule Information on Public Cloud for Preserving privacy

Tatsuya Miyagami[1], Atsushi Kanai[1], Noriaki Saito[2], Shigeaki Tanimoto[3], Hiroyuki Sato[4]

[1] Graduate School of Engineering Hosei University, Tokyo, Japan
tatsuya.miyagami.9t@stu.hosei.ac.jp, yoikana@hosei.ac.jp
[2] NTT Information Platform Laboratories, Tokyo, Japan
saito.noriaki@lab.ntt.co.jp
[3] Chiba Institute of Technology, Chiba, Japan
shigeaki.tanimoto@it-chiba.ac.jp
[4] The University of Tokyo, Tokyo, Japan
schuko@satolab.itc.u-tokyo.ac.jp

***Abstract*— We are currently experiencing an explosion of cloud technologies. However, a cloud service administrator may be an untrustworthy third party. Therefore, companies dealing with confidential information cannot use a public cloud. In this paper, we propose a method for preventing the leakage of private information on a cloud schedule service. In this method, even a cloud administrator or a hacker who steals a cloud service login key cannot read the true schedule because the schedule date is altered and schedule content is encrypted. Consequently we can safely use a public cloud schedule service with this method. We also evaluated the method's performance using an actual alteration program on Google server. We implement the proposed method and show the performance is practical by evaluating actually.**

Keywords-cloud computing; Internet security; privacy; Google Calendar; schedule service; date alteration.

I. INTRODUCTION

We are currently experiencing an explosion of cloud technologies. However, by considering "cloud" as a social infrastructure, we must also consider security [1]. For example, a cloud system is not managed by a user; therefore, cloud users cannot be certain that their critical information is actually safe [2].

Since schedule services are useful for a variety of fields, many people use them to manage their personal scheduling information. However, companies are unable to use scheduling services because information may include private or confidential information [3]. If there is a malicious administrator in the cloud, he might steal a user's privacy information or confidential information. If a malicious administrator exploits a company's confidential information, such as an important meeting schedule or customer information, its finances and reputation may be seriously damaged [4].

To solve this problem, it is necessary to protect private or confidential information from third party tapping. It is easy to preserve privacy or confidential information by encryption with respect to documents. The contents of a schedule can be encrypted on the schedule server. However, the schedule dates cannot be encrypted on the schedule server because if the value of the encrypted data becomes binary, the value cannot be saved in the schedule server as a schedule date. Therefore, the encryption of schedule dates cannot be used with a calendar service without changing the calendar interface.

In this paper, we propose a method for altering dates. The original schedule dates are completely changed to different dates. By using both alteration and encryption, a schedule can be protected from a third party. This is a good solution for managing both security and convenience of existing cloud schedule services at the same time. Note that we are not concerned here with encryption of schedule contents, which is rather simple, but with the alteration of schedule dates.

The organization of this paper is as follows. Section II describes related work. Section III introduces our alteration method for schedule services. Section IV describes the saving and reading algorithms created with our method. Section V discusses the evaluation of our alteration method. Section VI summarizes this paper.

II. RELATED WORK

Techniques of preserving privacy have been discussed for On-Line Analytical Processing (OLAP). Furthermore, Database-As-a-Service (DAS), which provides data management services for cloud computing, has become familiar.

 ISBN: 978-1-61208-176-2

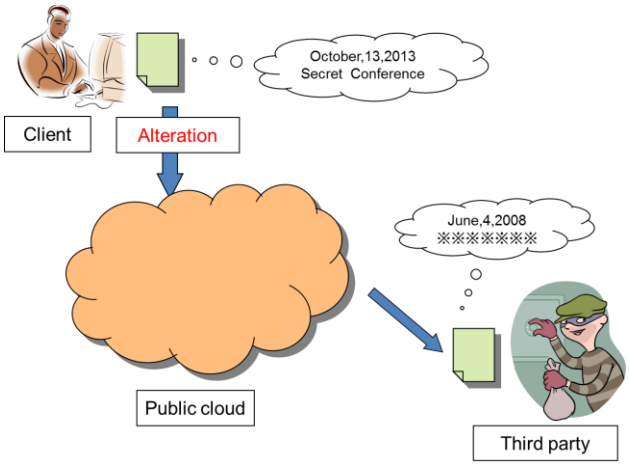


Figure 1. Alteration of schedule date in public cloud

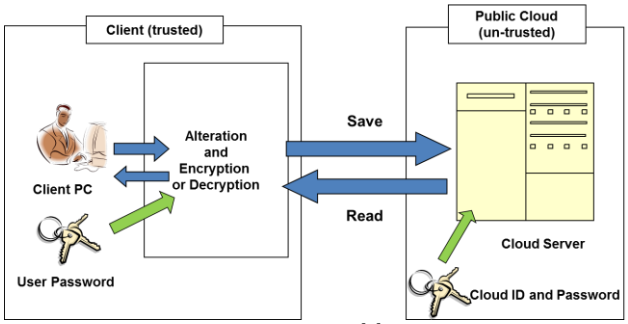


Figure 2. Trust model

With these technologies, server administrators may be untrustworthy third parties; therefore, privacy-preserving technologies have been necessary. In OLAP, perturbation techniques have been investigated for privacy-preserving data mining [5] [6] [7]. Using this technique, original values are perturbed and stored in a database; however, results of statistical queries remain correct. Consequently, privacy as original values is preserved. In DAS, cryptography has been commonly used to perform queries on encrypted data stored on a database [1] [8] [9].

The above techniques need to be applied to the basic functions of database systems, and it is necessary to replace or develop a new server system to use these technologies. On the other hand, we assume a current schedule server in the cloud. In this case, data types not treated on the schedule server cannot be used. This means dates need to be stored as dates in the schedule server through APIs. Therefore, dates cannot be encrypted because the binary value as an encrypted result cannot be stored in the date field in schedule databases. Furthermore, an altered date must be able to be decoded back to the original date but we need to maintain data mining results to be proper. For this reason, we developed a date alteration method.

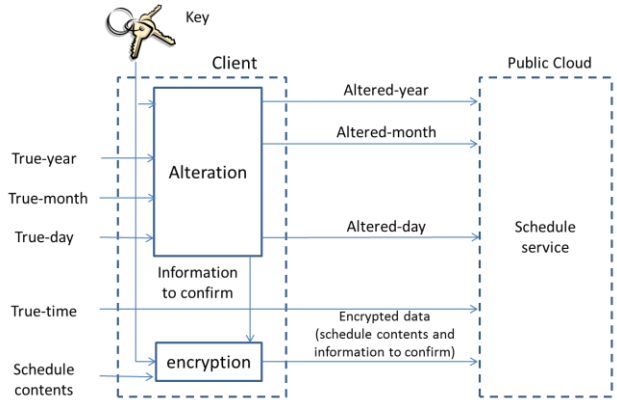


Figure 3. Saving schedule

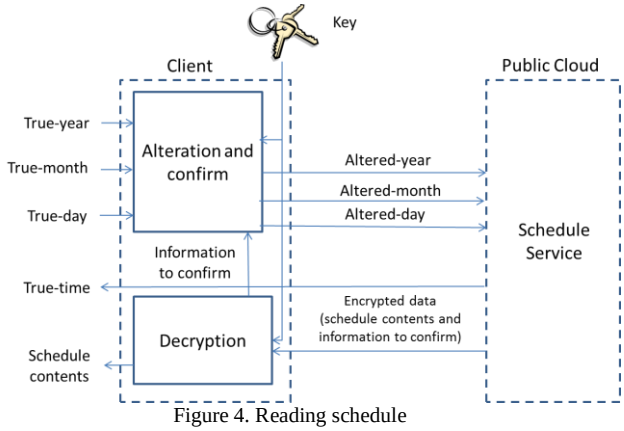


Figure 4. Reading schedule

III. CONCEPT OF ALTERATION METHOD OF SCHEDULE DATES

In this section, we describe the methodology and reading algorithms for altering dates.

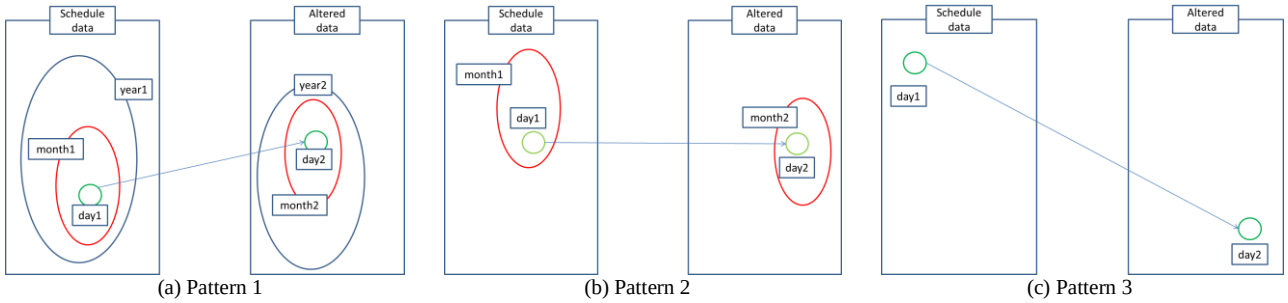
A. *Overview of alteration method*

The process of altering schedule dates is shown in Figure 1. First, the original schedule date is prepared on the client side. Next, the original schedule is converted to a different date by using a password, which is inputted and stored on the client side. Finally, the altered date is transferred to the cloud server.

B. *Trust model*

A trust model is shown in Figure 2. We assume a public cloud is not trusted; consequently, a password of a public cloud service for account authentication is also not trusted. For example, a security aware cloud [10][11] has been proposed with this kind of trust model. For this reason, another password (Key), which is different from the original password, and alteration of the original schedule date have to be prepared. This password must be kept on the client side, and there must be alteration and encryption modules on the client PC because the PC is assumed as trusted.

 ISBN: 978-1-61208-176-2



(a) Pattern 1 (b) Pattern 2 (c) Pattern 3

Figure 5. Three patterns of schedule date alteration

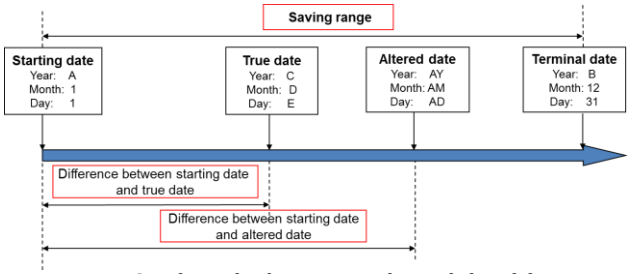


Figure 6. Relationship between true date and altered date

C. *Alteration*

A schematic of saving a schedule is shown in Figure 3. First, the original schedule data is divided into the elements of "true-year (TY)", "true-month (TM)", "true-day (TD)", "true-time", and "schedule-content" (which is the scheduled event). When the original date is altered, a Key, which is prepared by the client, is used. The TY, TM and TD are converted into "altered-year (AY)", "altered-month (AM)" and "altered-day (AD)" by using an alteration algorithm and the Key. We only altered the TY, TM, and TD and true-time was not. Note that when the true date is altered, the information necessary to confirm the true date is created. A detailed explanation of this information is described in Section IV.

A schematic of reading a schedule is shown in Figure 4. It is assumed that the TY, TM, and TD are known by the client, and true-time and schedule-content is unknown. The TY, TM and TD are altered again using the same Key. A client accesses the altered date, the encrypted schedule contents, and true-time. The schedule is decrypted with a Key, and the schedule contents and information to confirm the true date are divided. Finally, the true date is output using the information for confirming it.

D. *Date alteration pattern*

We divided our date alteration method into three patterns, and examined the efficiency of saving and reading a schedule. These three patterns are shown in Figure 5.

In Pattern 1, the day is converted to another within the same month, the month is converted to another within the same year, and the year is converted to another in the entire range of years contained in the system. In Pattern 2, the day is converted to another within the same month and the month is converted to another in the entire range of months contained in the system. In Pattern 3, the day is converted into another day within the entire range contained in the system.

The performance and degree of vulnerability differ depending on each pattern. This is discussed in more detail in Section V.

IV. SAVING AND READIG ALGORITHMS

In this section, we describe the saving and reading algorithms of dates.

These algorithms were created using our alteration method. The relationship between true date and altered date is shown in Figure 6. The user's password is input and converted into a series of numbers. These numbers are defined as a Key. The length of the Key should be 16 bits when using the exclusive-OR function. Moreover, when using a block cipher, the length of the key depends on the block cipher algorithm. Both algorithms are described as follows.

A. *Saving algorithm*

The saving algorithm of our alteration method is described as follows. The service's range means the entire period of the calendar in the specific service. Note that "A", "B", "C", "D", and "E" are defined as "year of starting date", "year of terminal date", "year of true date", "month of true date", and "day of true date".

(1) The difference between the true date and starting date of a particular service's range is calculated.

a) Pattern 1

$$dif\_year = C - A \quad (1)$$

b) Pattern 2

$$dif\_month2 = (C - A) \times 12 + (D - 1) \quad (2)$$

c) Pattern 3

The number of days from the starting date to schedule date are calculated using the function $F_1(x)$. Here, $F_1(x)$ is the number of days.

$$dif\_day3 = F_1(A,1,1,C,D,E) \quad (3)$$

(2) The temporal values are calculated using the exclusive-OR function or a block cipher.

a) Pattern 1

$$EY = dif\_year \oplus Key \quad (4)$$

$$EM = (D-1) \oplus Key \quad (5)$$

$$ED = (E-1) \oplus Key \quad (6)$$

The following equations are used if a block cipher is used.

$$EY = E_{Key}(dif\_year) \quad (7)$$

$$EM = E_{Key}(D-1) \quad (8)$$

$$ED = E_{Key}(E-1) \quad (9)$$

b) Pattern 2

$$EM2 = dif\_month2 \oplus Key \quad (10)$$

$$ED2 = (E-1) \oplus Key \quad (11)$$

The following equations are used if a block cipher is used.

$$EM2 = E_{Key}(dif\_month2) \quad (12)$$

$$ED2 = E_{Key}(E-1) \quad (13)$$

c) Pattern 3

$$ED3 = dif\_day3 \oplus Key \quad (14)$$

The following equations are used if a block cipher is used.

$$ED3 = E_{Key}(dif\_day3) \quad (15)$$

(3) The calculated values in A-(2) are calculated using the "modulo function" to interpose between the service's ranges. Note that the function is to provide the remainder of the division. The value of reminders using the mod function is the altered date. The value of the quotient using the "division function" is the information to confirm the true date mentioned in Section II-C. Note that the function is to provide the quotient of the division. This information is called the Element of Read Data (ERD). ND means the number of days in one month.

a) Pattern 1

A quotient is calculated with the altered date as a result of the respective divisions. The altered date is then determined to be AY1, AM1, and AD1.

$$AY1 = \mathrm{mod}(EY, B-A) + A \quad (16)$$

$$AM1 = \mathrm{mod}(EM, 12) \quad (17)$$

$$AD1 = \mathrm{mod}(ED, ND) \quad (18)$$

$$ERD_{year} = div(ED, B-A) \quad (19)$$

$$ERD_{month} = div(EM, 12) \quad (20)$$

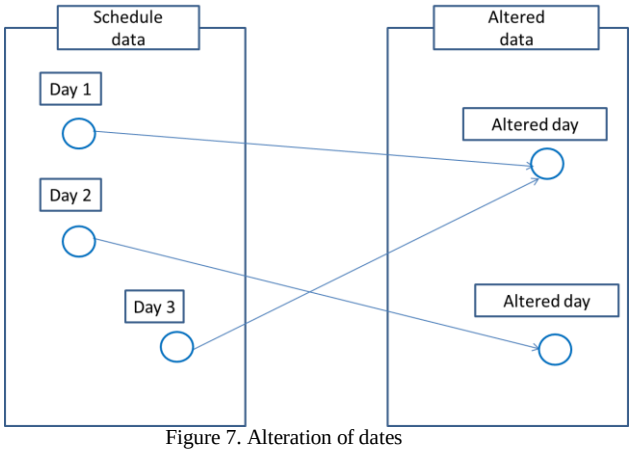


Figure 7. Alteration of dates

$$ERD_{day} = div(ED, ND) \quad (21)$$

b) Pattern 2

The year and month are combined and calculated as the number of months. SRM is the number of months in a service's range.

$$SRM = (B-A) \times 12 + 12 \quad (22)$$

$$AMN = \mathrm{mod}(EM2, SRM) \quad (23)$$

$$AD2 = \mathrm{mod}(ED2, ND) \quad (24)$$

$$ERD_{month\_number} = div(EM2, E-B) \quad (25)$$

$$ERD_{day2} = div(ED2, ND) \quad (26)$$

The altered date is determined to be AY2, AM2, and AD2.

$$AY2 = div(AMN, 12) + A \quad (27)$$

$$AM2 = \mathrm{mod}(AMN, 12) + 1 \quad (28)$$

c) Pattern 3

SRD is the number of days in a service's range. The days of the system range using $F_1(x)$ is calculated.

$$SRD = F_1(A,1,1,B,12,31) \quad (29)$$

$$ADN = \mathrm{mod}(ED3, SRD) + 1 \quad (30)$$

$$ERD_{day\_number} = div(ED3, SRD) \quad (31)$$

Using $F_2(x)$, the altered date is determined as AY3, AM3, and AD3. Here, $F_2(x)$ provides the altered-year, altered-month, and altered-day.

$$(AY3, AM3, AD3) = F_2(A,1,1,ADN) \quad (32)$$

(4) The contents of the schedule and ERD are concatenate.

(5) The above data is saved to the cloud server as the schedule contents. If necessary, the schedule contents are encrypted using the Key.

 ISBN: 978-1-61208-176-2

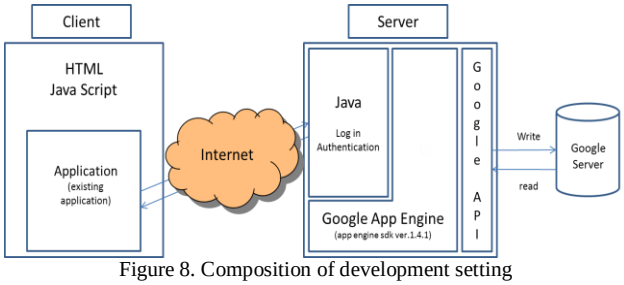


Figure 8. Composition of development setting

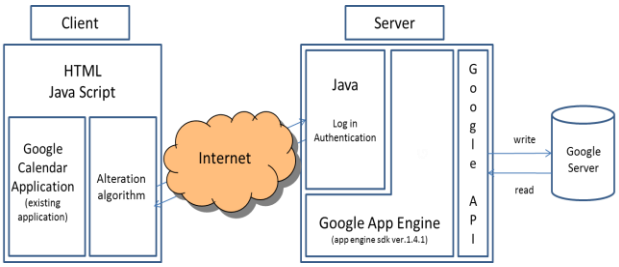


Figure 9. Final environment

B. *Reading algorithm*

The reading algorithm of our alteration method is described as follows.

(1) The altered date is calculated using the requested schedule date, and the saving algorithm is used to obtain the altered date.

(2) The schedule contents are read from the altered schedule date using the calendar interface.

(3) The ERD is extracted from the schedule contents and a true date is calculated. It is necessary to check the date because a true date can be altered many schedule data to one altered data, as shown in Figure 7. If this calculated date is equal to the requested date, the schedule contents become available; otherwise, they are rejected.

a) Pattern 1

The reverse order of the saving algorithm described in Section III-A is performed. The original date is calculated from the ERD.

$$TY = ((ERD_{year} \times (B - A) + AY) \oplus Key) + A \quad (33)$$

$$TM = ((ERD_{month} \times 12 + AM) \oplus Key) + 1 \quad (34)$$

$$TD = ((ERD_{day} \times ND + AD) \oplus Key) + 1 \quad (35)$$

When the block cipher is used, the true date is calculated as follows.

$$TY = (E_{Key}(ERD_{year} \times (B - A) + AY)) + A \quad (36)$$

$$TM = (E_{Key}(ERD_{month} \times 12 + AM)) + 1 \quad (37)$$

$$TD = (E_{Key}(ERD_{day} \times ND + AD)) + 1 \quad (38)$$

b) Pattern 2

The number of months from the starting date is computed using the altered year and altered month.

$$AMN = AY2 \times 12 + AM2 \quad (39)$$

The original date is calculated from the ERD.

$$TMN = ((ERD_{month\_number} \times SRM + AMN) \oplus Key) \quad (40)$$

$$TY2 = div(TMN, 12) + A \quad (41)$$

$$TM2 = \mathrm{mod}(TMN, 12) + 1 \quad (42)$$

$$TD2 = ((ERD_{day2} \times ND + AD2) \oplus Key) + 1 \quad (43)$$

When the block cipher is used, the true date is calculated as follows.

$$TMN = (E_{Key}(ERD_{month\_number} \times SRM + AMN)) \quad (44)$$

$$TD2 = (E_{Key}(ERD_{day2} \times ND + AD2)) + 1 \quad (45)$$

Note that if "(44)" is applied, "(41)" and "(42)" are applied, and TY2, TM2 are calculated.

c) Pattern 3

The number of days from the starting date is calculated using the AY and AM, and the difference in the days from starting date to the altered schedule date is calculated using $F_1(x)$.

$$ADN = F_1(A,1,1, AY3, AM3, AD3) \quad (46)$$

The true date is calculated from the ERD.

$$(TY3, TM3, TD3) = F_2(A,1,1, ((ERD_{day\_number} \times SRD + ADN) \oplus Key)) \quad (47)$$

When the block cipher is used, the original date is calculated as follows.

$$(TY3, TM3, TD3) = F_2(A, B, C, (E_{Key}(ERD_{day\_number} \times SRD + ADN))) \quad (48)$$

(4) The computed TY, TM, and TD are compared with C, D, and E. If the two values are equal, the calculated schedule becomes available.

V. EVALUATION AND DISCUSSION

A. *Implementation*

We developed alteration modules on "Google Calendar" [12], which is a type of software as a service (SaaS). We did not use an existing calendar application, but we used the Google Calendar API [13] to evaluate the algorithm. The environment of this module is shown as Figure 8.

 ISBN: 978-1-61208-176-2

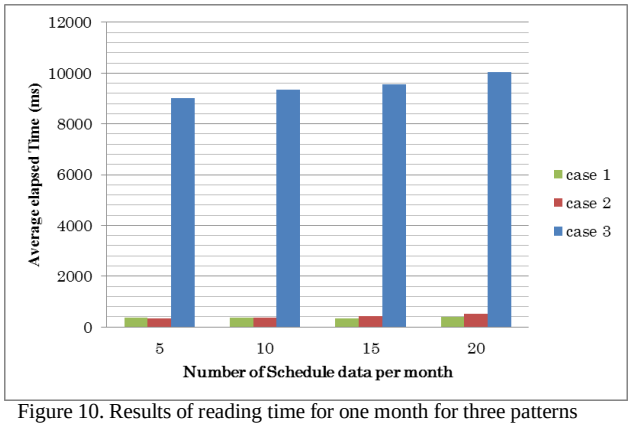


Figure 10. Results of reading time for one month for three patterns

TABLE I. RESULTS OF READING TIME FOR ONE MONTH FOR THREE PATTERNS

Number of schedule dates per month	*Pattern 1 (ms)*	*Pattern 2 (ms)*	*Pattern 3 (ms)*
5	353	334	9000
10	370	359	9333
15	340	418	9590
20	409	523	10049

In actual use, Google Calendar applications should be used for building the proposed algorithm into the application for user convenience, as shown in Figure 9.

B. Performance Evaluation

We evaluated the performance of the proposed method under the three patterns mentioned in the previous section, using the implemented modules discussed in the previous subsection.

B-1) Reading time for a month in all three cases.

B-2) Reading time for a year in Patterns 1 and 2.

We did not evaluate the performance of saving schedule data because there was no difference in performance among the three patterns. Therefore, we only evaluated performance of reading schedule data. When the schedule data is read, it is usually appropriate to read the data for one month or one week. In other words, it is not practical to use reading data for one year. However, by comparing the performances of the algorithm, the schedule data was read for a year.

For B-1, all schedules within the specified month given by a user are read, and the elapsed time of reading the schedules was evaluated in the three alteration patterns. The measured results for B-1 are shown in Figure 10. The elapsed time of Pattern 3 was much longer than those of the other two patterns. Patterns 1 and 2 produce the schedule for a month when making only one API call. On the other hand, Pattern 3 produce the schedule for a month by making an API call which is based on the number of days in a month.

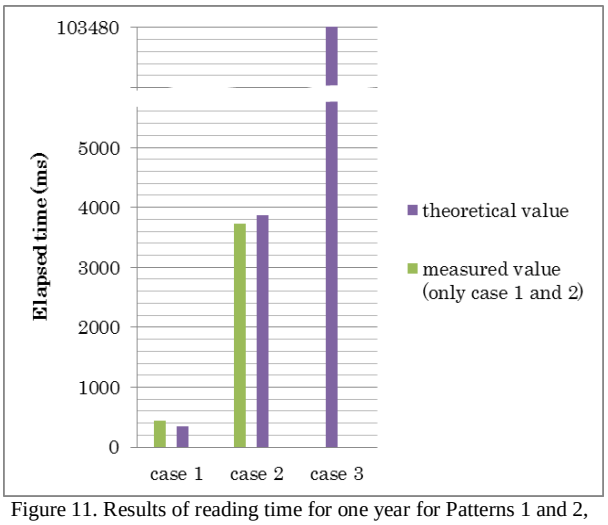


Figure 11. Results of reading time for one year for Patterns 1 and 2, and calculated theoritical value for Pattern 3

TABLE II. RESULTS OF READING TIME FOR ONE YEAR FOR PATTERN 1 AND 2

Number of schedule dates Per month	*pattern 1 (ms)*	*pattern 2 (ms)*
20	442	3731

There were not many differences in the elapsed time for Pattern 3 when the number of schedule dates included in one month increased. As a result, the elapsed time did not much depend on the number of schedule dates in B-1 because the number of API calls increased during the elapsed time.

Equation "(49)" is derived from Figure 10. Note that "T" is the elapsed time.

$$T = t_n \times (A - x) + t_e \times x \quad (49)$$

t_n : The elapsed time when there are no date for one API call.

t_e : The elapsed time when there is a date for one API call.

x : The number of API calls

A : The maximum number of API calls

Here, t_n and t_e are calculated from the elapsed time, which is measured for a month (maximum number of days is 31). Note that t_n and t_e are assumed to be constant numbers.

For calling schedule dates of a month, we substitute $A = 31$, $x = 5$, and $T = 9000$ ms into "(49)".

$$t_n \times (31 - 5) + t_e \times 5 = 9000 \quad (50)$$

When $x = 10$,

$$t_n \times (31 - 10) + t_e \times 10 = 9300 \quad (51)$$

From "(50)" and "(51)", t_n and t_e are calculated as

$$t_n = 280\,(\text{ms}) \quad (52)$$

$$t_e = 344\ (\text{ms}) \quad (53)$$

By substituting "(52)" and "(53)" into "(49)", T is represented as

$$T = 280 \times (A - x) + 344 \times x \quad (54)$$

When substituting $x = 20$ to confirm the elapsed time,

$$T = 280 \times (31 - 20) + 344 \times 20 = 9960\,(\text{ms}) \quad (55)$$

The calculated value was close to the measured value.

For B-2, we compared the elapsed time of reading schedule data for a year. The measured results for B-2 are shown in Figure 11. The elapsed time of Pattern 2 was longer than that of Pattern 1. This pattern obtained the schedule for a year with one API call. On the other hand, Pattern 2 obtained the schedule for a year with twenty API calls. Only the theoretical value of Pattern 3 was published in Figure 11 at this time because this pattern was not able to make API calls for a year in the experimental environment.

The theoretical value of the elapsed time was calculated. Using "(54)".

The elapsed time of pattern 1 is

$$T = 280 \times (1 - 1) + 344 \times 1 = 344\,(\text{ms}) \quad (56)$$

The elapsed time of pattern 2 is

$$T = 280 \times (12 - 8) + 344 \times 8 = 3872\,(\text{ms}) \quad (57)$$

The elapsed time of pattern 3 is

$$T = 280 \times (365 - 20) + 344 \times 20 = 103480\ (\text{ms}) \quad (58)$$

When "(56)", "(57)" and Table II were compared, the actual measurement and theoretical values were close.

According to Tables I and II, the elapsed time was at most 10 seconds. In a cloud environment, it is thought that the processing time increases. Therefore, an elapsed time of 10 seconds is appropriate for practical use. However, since the theoretical value of Pattern 3 was at most about 103 seconds, it is not practical for reading schedule dates for a year.

The elapsed time is proportional to the number of API calls. Therefore, it can be said that Pattern 1 with API calls is superior to the others patterns in terms of performance. Therefore, the degree of module performance is higher in reverse order, Pattern 1 > Pattern 2 > Pattern 3.

To improve the elapsed time, it is necessary to develop an algorithm for reducing the number of API calls.

C. *Security of alteration*

A possible attack is described as follows.

First, the original schedule date may be predicted by a third party. As mentioned in Section II, Pattern 3 is the safest pattern because a date is mapped throughout the schedule range. Alteration of Pattern 2 is performed day to day within a month. Therefore, altered date distribution does not change month to month compared to the true distribution. Therefore, if a hacker observes a newly added schedule, he may know what month the current month is because the current month must be the month that the number of added schedules is largest. Therefore, it is difficult to guess the true date.

A user prepares a password (Key) beforehand for altering the date. However, if the same key is used for a long time, the schedule date range will not be large; therefore, a hacker can predict the key by brute force. To prevent this kind of attack, a user should periodically change the key. When the key is changed, it is necessary to simultaneously calculate the ERD again.

Second, the original schedule might be guessed from the altered schedule. A schematic of guessing the true schedule from an altered one is shown in Figure 12. For instance, there are many schedules in 2011, and there were few in 2010. First, 2010 and 2011 are altered to 2030 and 2005, respectively. It is assumed that the server administrator knows the schedule frequency in 2010 and 2011 of a client. The server administrator investigates the altered schedule frequency in 2030 and 2005 and compares each year. If it turns out that 2005 had many schedules, it may turn out the 2005's altered schedules are equal to those of 2011. This is the same not only for the combination of "year and month" but also "month and day." Therefore, Pattern 3 is safest, and Pattern 2 is safer than Pattern 1.

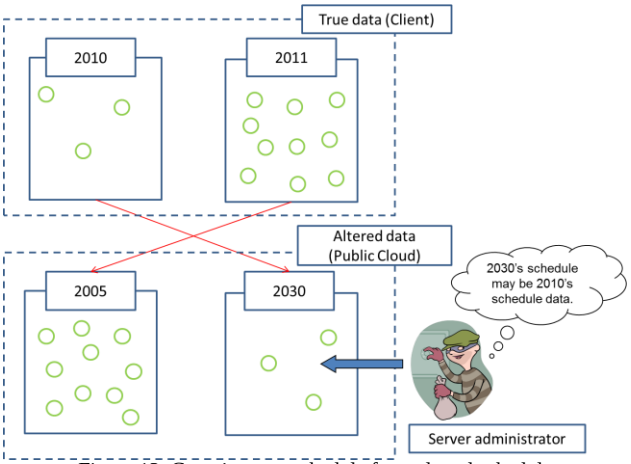


Figure 12. Guessing true schedule from altered schedule

TABLE III. AVAILABLE FUNCTIONS AND APIS

Function of current schedule APIs	*Availability of alteration*
Create an event	○
Create a new calendar	○
Repeat an event	○
Privacy settings for individual events	○
Edit or view event details	○
Delete or remove an event	○
Delete a calendar	○
Events that last all day	○
Color Code an Event	○
Edit your calendar name	○
Notifications (Daily Agenda)	×
Notifications (Event Reminders)	×

D. Available APIs

Schedule dates are all altered on the Google Calendar server, so some APIs might not perform properly. Existing Google Calendar APIs are listed in Table III. Existing schedule services have many functions. For instance, a notification API will be set at the date after alteration. As a result, an alarm will be activated on the wrong date.

VI. CONCLUSION AND REMARKS

There are various advantages in cloud computing; however, there are still many security problems. We proposed an alteration method to protect private or confidential information from third party tapping. We also implemented two alteration modules and evaluated the proposed method's performance in three patterns on Google calendar. We found that the method's performance was lower than usage without alteration, but it is still useful. We plan to develop a better performing algorithm in the future.

For actual use, Google Calendar applications should be used with the proposed modules built into the application for user convenience. This is for future work.

REFERENCES

[1] C. Almond, "A Practical Guide to Cloud Computing Security", Accenture and Microsoft, August 27, 2009

[2] "Public Cloud Computing Security Issues". [Online] Available:http://www.thebunker.net/managed-hosting/cloud/public-cloud-computing-security-issues/, <retrieved: November, 2011>

[3] D. Yuefa, W. Bo, G. Yaqiang, Z. Quan, and T. Chaojing, "Data Security Model for Cloud Computing", ISBN 978-952-5726-06-0, Proceeding of the 2009 International Workshop on Information Security and Application (IWISA 2009) Qingdao, China, November 21-22, 2009

[4] BalaBit IT security, "Cloud Security Risks and Solutions", First Edition, July 1, 2010

[5] R. Agrawal, R. Strikant, and D. Thomas, "Privacy Preserving OLAP" Proc. 25th ACM SIGMOD Int'1 Conf. Management of Data, ACM Press, 2005, pp. 251-262

[6] N. Zbang and W. Zbao, "Privacy-Preserving Data Mining Systems", IEE Computer Society Computer, April 2007, pp. 52-58

[7] J. Vaidya and C. Clifton, "Privacy-Preserving Data Mining: Why, How, and When", IEEE Security&Privacy Building Confidence in a Networked World, November/December, 2004

[8] D. Boneh, G. Crescenzo, R. Ostrovsky, and G. Persiano, "Public Key Encryption with Keyword Search", Proceedings of EUROCRYPT '04, vol. 3027

[9] Z. Yang, S. Zhong, and N. Wright, "Privacy-Preserving Queries on Encrypted Data", Proceedings of the 11th European Symposium On Research In Computer Security (Esorics), LNCS4189, pp. 479-495, 2006.

[10] H. Sato, A. Kanai, and S. Tanimoto, "A Cloud Trust Model in Security Aware Cloud", Proceedings of 10th International Symposium on Applications and the Internet (SAINT 2010), pp. 121

[11] H. Sato, A. Kanai, and S. Tanimoto, "Bulding a Security Aware Cloud by Extending Internal Control to Cloud", Proceedings of 10th International Symposium on Autonomous Decentrakized Systems (ISADS 2011), 2011.

[12] Google, "Date API Developer's Guide".[Online] Available:http://code.google.com/intl/ja/apis/calendar/data/2.0/developers_guide.html <retrieved: November, 2011>

[13] Google, "Calendar help". [Online] Available: http://www.google.com/support/calendar/ <retrieved: November, 2011>

 ISBN: 978-1-61208-176-2

Identifying Potentially Useful Email Header Features for Email Spam Filtering

Omar Al-Jarrah*, Ismail Khater‡ and Basheer Al-Duwairi†
*Department of Computer Engineering
†Department of Network Engineering & Security
Jordan University of Science & Technology, Irbid, Jordan 22110
‡Department of Computer Systems Engineering
Birzeit University, Birzeit, West Bank, Palestine
Email: aljarrah@just.edu.jo, ikhater@birzeit.edu, basheer@just.edu.jo

***Abstract*—Email spam continues to be a major problem in the Internet. With the spread of malware combined with the power of botnets, spammers are now able to launch large scale spam campaigns causing major traffic increase and leading to enormous economical loss. In this paper, we identify potentially useful email header features for email spam filtering by analyzing publicly available datasets. Then, we use these features as input to several machine learning-based classifiers and compare their performance in filtering email spam. These classifiers are: C4.5 Decision Tree (DT), Support Vector Machine (SVM), Multilayer Perception (MP), Nave Bays (NB), Bayesian Network (BN), and Random Forest (RF). Experimental studies based on publicly available datasets show that RF classifier has the best performance with an average accuracy, precision, recall, F-Measure, ROC area of 98.5%, 98.4%, 98.5%, and 98.5%, respectively.**

***Index Terms*—Email Spam, Machine Learning**

I. Introduction

Email spam, defined as unsolicited bulk email, continues to be a major problem in the Internet. Spammers are now able to launch large scale spam campaigns, malware and botnets helped spammers to spread spam widely. Email spam cause many problems, increase traffic and leading to enormous economical loss. Recent studies [1], [2] revealed that spam traffic constitute more than 89% of Internet traffic. According to Symantec [3], in March 2011 the global Spam rate was 79.3%. The cost of managing spam is huge compared with the cost of sending spam which is negligible. It includes the waste of network resources and network storage, the cost of traffic and the congestion over the network, in addition to the cost associated with the waste in employees' productivity. It was estimated that an employee spends 10 minutes a day on average sorting through unsolicited messages [4]. Other studies [5], [6], [7] reported that spam costs billions of dollars. Ferris Research Analyzer Information Services estimated the total worldwide financial losses caused by spam in 2009 as $130 billion; $42 billion in the U.S. alone [8].

Spammers are increasingly employing sophisticated methods to spread their spam emails. In addition, they employ advanced techniques to evade spam detection. A typical spam campaign involves using thousands of spam agents to send spam to a targeted list of recipients. In such campaigns, standard spam templates are used as the base for all email messages. However, each spam agent substitutes different set of attributes to obtain messages that do not look similar. Moreover, spammers are increasingly adopting image-based spam wherein the body of the spam email is converted to an image, which renders text-based and statistical spam filters useless.

Blocking spam email is considered a priority for network administrators and security researchers. There have been tremendous research efforts in this field that resulted in a lot of commercial spam filtering products. Header-based email spam filtering is considered as one of the main approaches in this field. In this approach, a machine learning classifier is applied on features extracted from email header information to distinguish ham from spam, and the accuracy of the header-based email spam filter depends greatly on the email header fields used for feature selection. In this paper, we identify potentially useful email header features based on analyzing large publicly available dtatsets to determine the most distinctive features. Also, we include most of the mandatory and optional email header fields in order to fill any gab or missing information that is required for email classification.

This paper presents a performance evaluation of several machine learning-based classifiers and compare their performance in filtering email spam based on email header information. It also proposes including important email header features for this purpose. The rest of this paper is organized as follows: Section II reviews related work. Section III discusses the main features of email header considered in our work. Section IV evaluates the performance of different machine learning-based classifiers in filtering header-based email spam. Finally, Section V concludes the paper.

II. Related work

An email message typically consists of header and body. The header is a necessary component of any email message. The Simple Mail Transfer Protocol (SMTP) [15] defines a set of fields to be contained in the email message header to achieve successful delivery of email messages and to provide important information for the recipient. These fields include:

 ISBN: 978-1-61208-176-2

email history, email date, time, sender of the email, receiver(s) of the email, email ID, email subject, etc. Header-based email spam filtering represents an efficient and lightweight approach to achieve filtering of spam messages by inspecting email message header information. Typically, a machine learning classifier is applied on features extracted from email header information to distinguish ham from spam. For example, Sheu [10] categorized emails into four categories based on the title: sexual, finance and job-hunting, marketing and advertising, and total category. Then he classified them according to the attributes from email message header. He proposed a new filtering method based on categorized Decision Tree (DT), namely, applying the Decision Tree technique for each of the categories based on attributes (features) extracted from the email header. The extracted features are from the sender field, email's title, sending date, and the email's size. Sheu applied his filter on a Chinese emails and obtained accuracy, precision, and recall of 96.5%, 96.67%, 96.3%, respectively.

Wu [11] proposed a rule-based processing that identifies and digitizes the spamming behaviors observed from the headers and syslogs of emails by comparing the most frequent header fields of these emails with their syslog at the server. Wu noticed the differences in the header filed of the sent email from what is recorded in the syslog, and he utilized that spamming behavior as features for describing emails. A rule-based processing and back-propagation neural networks were applied on the extracted features. He achieved an accuracy of 99.6% with ham misclassification of 0.63%. Ye et al. [12] proposed a spam discrimination model based on SVM to sort out emails according to the features of email headers. The extracted features from email header fields are the return-path, received, message-id, from, to, date and x-mailer; They used the SVM classifier to achieve a recall ratio of 96.9%, a precision ratio of 99.28%, and an accuracy ratio of 98.1%.

Wang [13], presented a statistical analysis of the header session message of junk and normal emails and the possibility of utilizing these messages to perform spam filtering. A statistical analysis was performed on the contents of $10,024$ junk emails collected from a spam archive database. The results demonstrated that up to 92.5% of junk emails are filtered out when utilizing mail user agent, message-id, sender and receiver addresses as features.

Recently, Hu et al. [9] proposed an intelligent hybrid spam-filtering framework to detect spam by analyzing only email headers. This framework is suitable for extremely large email servers because of its scalability and efficiency. Their filter can be deployed alone or in conjunction with other filters. The extracted features from the email header are the originator field, destination field, x-mailer field, sender server IP address, and email subject. Five popular classifiers were applied on the extracted features: Random Forest (RF), C4.5 Decision Tree (DT), Nave Bayes (NB), Bayesian Network (BN), and Support Vector Machine (SVM). The best performance was obtained by the RF classifier with accuracy, precision, recall , and F-measure of 96.7%, 92.99%, 92.99%, 93.3%, respectively. These results were obtained when applying the classifiers on a dataset of $33,209$ emails and another dataset of 21,725 emails. The work presented in this paper focuses mainly on potentially useful header features for email spam filtering. These features were selected by analyzing publicly available dtatsets (described in Subsection IV-B). Table I provides a summary of the main email header features considered by different spam filtering techniques as reported in the literature. It also shows the main features that we consider in our work.

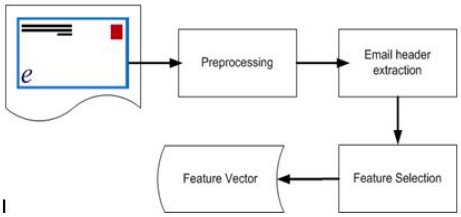


Fig. 1. The process of building feature vector of an email

III. Feature Selection

Feature selection represents the most important step of Header-based email spam filtering technique. In this step, we study information available in the email message header and carefully select some of them to be among the features used for classification. It is important to mention that the selection of email header features is based on analyzing large publicly available dtatsets (described in Subsection IV-B) to determine the most distinctive features. It is also important to point out that we include most of the mandatory and optional email header fields in order to fill any gab or missing information that is required for email classification. Figure 1 shows the process of building a feature vector of an email. This process starts by preprocessing of email messages to convert them into a standard format as described in RFC 2822. After that, we extract the header of the email to select the required features and build the feature vector which summarizes all the needed information from an email. This feature vector is then used to build the feature space for all emails that are needed for the classification phase.

The following subsections describe the fields of email message header that we consider in our work which turn to be of important value to classify email messages.

A. Received Field

Each email can contain more than one "Received" field. This field is typically used for email tracking by reading it from bottom to top. The bottom represents the first mail server that got involved in transporting the message, and the top represents the most recent one, where each received line represents a handoff between machines. Hence, a new received field will be added on the top of the stack for each host received the email and transport it, and to which host the message will be delivered, in addition to the time and date of passing. The following are the features that we extract from this field:

 ISBN: 978-1-61208-176-2

TABLE I
EMAIL HEADER FEATURES CONSIDERED BY DIFFERENT MACHINE LEARNING SPAM FILTERING TECHNIQUES

Sheu, 2009 [10]	Ye et. al., 2008 [12]	Wu, 2009 [11]	Hu et. al., 2010 [9]	Wang & Chen, 2007 [13]	Our Approach
Length of sender field , Sender field, Title (more than one category) , Time, Size of email	Received field (domain add., IP add., relay servers, date, time), From field, To field, Date field, Message-ID, X-Mailer	Comparing header fields with syslog	Originator fields, Destination fields, X-Mailer field, Sender IP, Email subject	Sender address validity, Receiver address (To, CC, BCC), Mail User Agent, Message-ID	Received field # of hops, Span Time, Domain add. Legality, Date & Time Legality, IP add. Legality, sender add. legality, # of Receivers (To, CC, BCC), Mail User Agent, Message-ID, Email subject Date of reception

1. *The number of hops*. This feature represents the number of the relay servers used to deliver the message from its origin to it its final destination. It was noticed based on different datasets that most of spam messages have a small number of hops. That means the spammers have exploited a predefined relay servers for delivering their spam, so the number of hops is limited, while in the normal case the number of relay servers may vary according to the paths the message follow to reach its final destination.
2. *Span time*. Span time represents the total time of the email through its journey from its origin to its final destination. This feature is considered as one of the most important features in our work. It is noticed that most of the spam emails have a large span time as compared to legitimate emails and some of them is negative in value.
3. *Domain address existence*. Domain address existence feature expresses whether the domain address of the host that delivers the message exists or not. This could be of little value to discriminate the spam emails from ham emails, but we keep it as a supporting feature.
4. *Date and time legality*. The purpose of this feature is to discover illegal date and time of email messages. The idea here is to check the date and time of email messages as they travel from one relay server to another. We believe this is an important feature because typically the date and time of legitimate email servers would be adjusted correctly. However, this is not necessarily the case for compromised machines that are used as email relays as we have discovered in the spam dataset.
5. *IP address legality*. This feature checks the legality of the host IP address, because spammers tend to hide or obfuscate IP addresses of their spam messages in order to avoid being blacklisted. We just check the format and the existence of the IP address.

B. Sender Address Legality

This feature is a conventional feature that is mentioned in most of the header based filters. The "From" field is one of the mandatory fields that every email must include, so the absence of this field is a cue for spamming behavior, the spammers tend to hide or use fake email addresses in order to avoid being blacklisted.

C. Number of Receivers

The recipients addresses of an email message are listed in one or more of the "To", "CC", and "BCC" fields. The "To" field contains the addresses of the primary recipients and the carbon copy "CC" field contains the addresses of the secondary recipients of the email, while the blind carbon copy "BCC" field contains the addresses of the recipients that are not included in copies of the email sent to the "To" and "CC" recipients. Many studies (e.g., [9], [13]) showed that spammers prefer to use the "BCC" field in order to send spam emails to a large number of recipients, at the same time no one of the recipients can obtain the list of the addresses that are collected by the spammers, because the SMTP server send a separate email to each one of the recipients listed in the "BCC" field, and every recipient has no information about the other recipients. In fact, most of the spam emails usually have small number of addresses in the "To" field which suggests that these emails were originally sent to many recipients using the "BCC" field such that individual recipients would not be able to identify other recipients of the same email.

D. Date of Reception

The "Date" field is a mandatory field that represents the date and time of the email when it is sent by the sender at the Mail User Agent (MUA). It is to be mentioned that the time recorded in this field is based on the location of the mail server of the sender which could belong to a time zone different from that of the recipient. Therefore, we convert all timing information into Universal Time Coordination (UTC) to have a common base for comparison. Basically, we compare the date of sending the email with the date of reception as recorded at the final hop in the "Received" field. We noticed that most spam emails do not have valid date of reception which suggests that this feature could be very helpful in our study.

E. Mail User Agent (MUA)

This is an optional field in the email header, appears as "X-Mailer" field which contains the email program used for the generation of the email. In this field, the email client or MUA name and version is recorded. Spammers usually tend to leave this field empty or fill it with random text. Based on

 ISBN: 978-1-61208-176-2

that, we take this field into consideration by checking whether it is existing or it is missing from the email message header.

F. Message-ID

This is a globally unique ID for each generated message. The "Message-ID" field is a machine readable ID which takes the name of the machine and the date and time of the email when it is sent. This field consists of two parts separated by @ sign. The right side part specifies the domain name or the machine name. This could be of a particular interest, because we noticed that most of spammers tend to hide this part or even fake the domain name to avoid being blacklisted. Therefore, it is required to make sure that the domain name in the "Message-ID" field is the same as the domain name in the "From" field. Inconsistency of this information would indicate a spamming behavior. It is important to mention here, that some mail user agents append the machine name to the domain name to the right of the @ sign. To overcome this issue, we used the partial matching with the domain name in the "From" field, and we noticed mismatches in most of the spam emails.

G. Email Subject

The subject contains a limited number of characters as described in RFC 822 and RFC 2822 [15]. It contains the topic and a summary of the email. Spammers may exploit the subject and use some special characters or words (e.g., "Try it for free!", "$ Money Maker $", "* URGENT ASSISTANT NEEDED *", etc.) to attract the user to open the email. Therefore, having special characters/phrases in the subject line may strongly indicate that the email is spam.

IV. Performance Evaluation

In this section, we evaluate the performance of several machine learning-based classifiers and compare their performance in filtering email spam based on email header information mentioned in Section III. In particular, we consider C4.5 Decision Tree (DT), Support Vector Machine (SVM), Multilayer Perception (MP), Nave Byays (NB), Bayesian Network (BN), and Random Forest (RF). Basically, our experiments involve evaluating the performance of these classifiers in terms of accuracy, precession, recall, and F-measure as defined Subsection IV-A using publicly available datasets. Email spam datatsets have been divided into a train and test sets according to the cross validation technique, where we used 10-fold cross validation. Weka tool [14] has been used for applying the machine learning techniques. Weka requires that the used features must conform to the input format of Weka. Therefore, the used features were ordered in a CSV file in the following format:

feature 1, feature 2, , feature n, class label

By default the class labels are located at the end of each row. In our experiments, we have two class labels used to categorize the image in the email, a legitimate email is marked as *Ham*, while the spam email is marked as *Spam*.

Prediction	Actual	
	Spam	Ham
Spam	TP	FN
Ham	FP	TN

Fig. 2. Confusion Matrix

A. Performance Metrics

We use the following standard performance metrics to evaluate the proposed technique: accuracy, precision, recall, F-measure, which are defined as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F - measure = \frac{2Precision.Recall}{Precision + Recall} \quad (4)$$

where *FP, FN, TP, TN* are defined as follows:

- *False Positive (FP):* The number of misclassified legitimate emails.
- *False Negative (FN):* The number of misclassified spam emails.
- *True Positive (TP):* The number of spam messages that are correctly classified.
- *True Negative (TN):* The number of legitimate emails that are correctly classified.

Precision is the percentage of correct prediction (for spam email), while spam Recall examines the probability of true positive examples being retrieved (completeness of the retrieval process), which means that there is no relation between precision and recall. On the other hand, F-measure combines these two metrics in one equation which can be interpreted as a weighted average of precision and recall. In addition, we use Receiver Operating Characteristics (ROC) curves which are commonly used to evaluate machine learning-based systems. These curves are basically a two-dimensional graphs where TP rate is plotted on y-axis and FP rate is plotted on x-axis. Therefore, depicting the tradeoffs between benefits TP and costs FP [19]. A common method to compare between classifiers is to calculate the Area Under ROC Curve (AUC).

It is important to mention that our definition of the performance metrics is mainly based on the confusion matrix shown in Figure 2.

B. Datasets

Our experiments are based on the following two publicly available recent datasets.

- CEAS2008 live spam challenge laboratory corpus [16] which contains 32703 labeled emails. Among these emails there are 26180 spam emails and 6523 ham

 ISBN: 978-1-61208-176-2

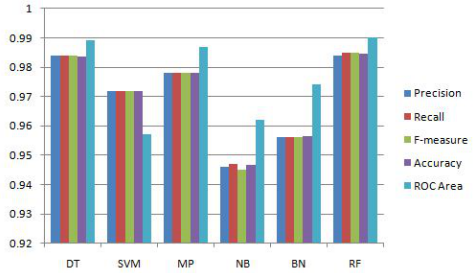


Fig. 3. The Performance of different machine learning techniques applied on CEAS2008 dataset in terms of Accuracy, precession, recall, F-measure, and ROC area

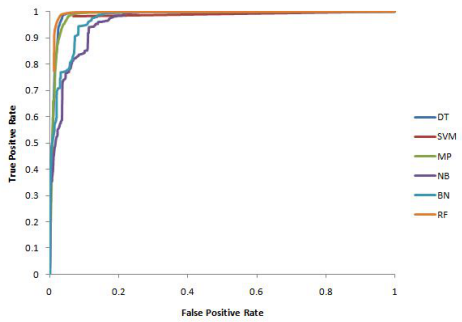


Fig. 4. ROC curves for the six classifiers applied on CEASE2008 dataset

email, this dataset was collected during the CEAS 2008 conference and it is considered as one of the TREC public spam corpus.

- CSDMC2010 spam corpus [18]. This dataset contains 4327 emails out of which there are 2949 non-spam (ham) emails and 1378 spam emails.

It is important to mention that these datasets were used for training and testing.

C. Experimental Results

1) Results based on CEAS2008 dataset: Figure 3 depicts the performance of the different classifiers in terms of accuracy, precision, recall, F-measure and the area under ROC. This figure shows the disparity among the classifiers in terms of precision, recall, F-measure and accuracy. It can be seen that RF classifier outperform all the other classifiers with an average accuracy, precision, recall, F-Measure, ROC area of 98.5%, 98.4%, 98.5%, 98.5%, and 99%, respectively. The ROC curves for all classifiers considered in this study are shown in Figure 4. This figure confirms that the RF classifier has the best performance compared to other classifiers as it maintains the best balance between false positive rate and true positive rate. DT classifier comes after RF classifier, then MP and SVM classifiers, while the BN and NB classifiers comes last. NB classifier was the worst in this group.

It is important to be mentioned that the results of other classifiers were as follows: DT classifier achieved an average precision and recall of 98.4%, which indicates that DT classifier succeeds in classifying most of the emails based on their header information. For the SVM classifier, it can be seen that it achieved good results for this dataset. However, the results were not that good in case of small size dataset as described in Subsection IV-C2. The other issue is the trade-off between FP and FN, which can be described by the ROC area. In the case of the MP classifier, the datasets were divided using the cross validation technique. Having the trained network; we can use it in recognizing spam emails of the testing set by invoking the simulation function, which takes the input feature vector and the trained network as inputs and computes the outputs according to the weights of the neurons, then it finds the output of the maximum weight. This classifier achieved an average precision and recall of 97.8%.

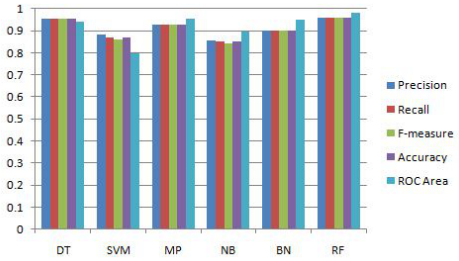


Fig. 5. The Performance of different machine learning techniques applied on CSDMC2010 dataset in terms of Accuracy, precession, recall, F-measure, and ROC area

2) Results Based on the CSDMC2010 dataset : In order to confirm the results obtained using CEAS2008 dataset, we repeated our experiments using another recent dataset (however, with smaller size). Figure 5 depicts the performance of the different classifiers using this dataset in terms of accuracy, precision, recall, F-measure and the area under ROC. It can be seen that RF classifier outperform all the other classifiers with an average accuracy, precision, recall, F-Measure, ROC area of 95.8%, 95.8%, 95.8%, 95.8% and 98.1%, respectively. It is to be noted that all classifiers achieved comparable performance this time indicating that the performance of some classifiers depends on the dataset used for testing and training.The MP classifier was very successful in recognizing 99In both cases, RF classifiers was on top of thlist in terms of performance. The ROC curves for all classifiers considered in this study are shown in Figure 6. This figure confirms that the RF classifier has the best performance compared to other classifiers as it maintains the best balance between false positive rate and true positive rate.

D. Comparison with Previous Work

In this subsection, we compare the performance of the proposed scheme with other header-based email spam filtering techniques ([9], [10], [11], [12], [13]) based on the results reported in the literature for these techniques. Table II shows the best performance obtained them and compare it to the results obtained using the proposed work. It can be seen that applying RF classifier to the email header features described in Section III results in better performance as compared to

TABLE II
PERFORMANCE OF THE PROPOSED WORK COMPARED TO OTHER HEADER-BASED EMAIL SPAM FILTERS. A: ACCURACY, P: PRECISION, R: RECALL, F: F-MEASURE

Spam Filter	Sheu, 2009 [10]	Ye et al., 2008 [12]	Wu, 2009 [11]	Hu et al., 2010 [9]	Wang & Chen, 2007 [13]	Our Approach
Classifier(s) used	DT	SVM	Rule-based & back-propagation NN	RF, DT, NB, BN, SVM	Statistical analysis	DT, SVM, MP, NB, BN, RF
Best performance obtained	A=96.5%, P=96.67%, R=96.3%	A=98.1%, P=99.28%, R=96.9%	A=99.6 % (ham misclassification = 0.63%)	RF (A=96.7%, P=93.5%, R=92.3%, F=93.3%)	92.5% of junk emails are filtered out	RF (A=98.5% , P=98.9%, R=99.2%, F=99%)

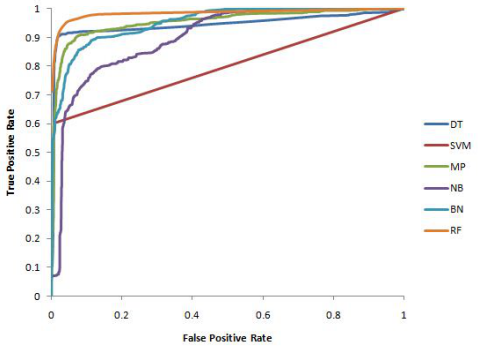


Fig. 6. ROC curves for the six classifiers applied on CSDMC2010 dataset

other header-based spam filters.

V. CONCLUSION

Spammers are increasingly employing sophisticated methods to spread their spam emails. Also, they employ advanced techniques to evade spam detection. A typical spam campaign involves using thousands of spam agents to send spam to a targeted list of recipients. In such campaigns, standard spam templates are used as the base of all email messages. However, each spam agent substitutes different set of attributes to obtain messages that do not look similar. In this paper, we evaluated the performance of several machine learning-based classifiers and compared their performance in filtering email spam based on email header information. These classifiers are: C4.5 Decision Tree (DT), Support Vector Machine (SVM), Multilayer Perception (MP), Nave Bays (NB), Bayesian Network (BN), and Random Forest (RF). We adopted header-based email spam filtering by including additional header information features that found to be of a great importance to improve the performance of this technique. We evaluate the proposed work through experimental studies based on publicly available datasets. Our studies show that RF classifier outperform all the other classifiers with an average accuracy, precision, recall, F-Measure, ROC area of 98.5%, 98.4%, 98.5%, 98.5%, and 99%, respectively.

REFERENCES

[1] C. Kreibichy, et al., "Spamcraft: An Inside Look At Spam Campaign Orchestration," Proceedings of the Second USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET '09), Boston, Massachusetts, April 2009.

[2] M. Intelligence, "MessageLabs Intelligence: 2010 Annual Security Report," 2010. Retrieved: July, 2011. Available at: http://www.clearnorthtech.com/images/MessageLabsIntelligence_2010_Annual_Report.pdf

[3] Symantec. March 2011 Intelligence Report. Retrieved: July, 2011. Available at: http://www.symantec.com/about/news/release/article.jspprid=20110329_01

[4] S. Hinde, "Spam, scams, chains, hoaxes and other junk mail," Computers & Security, vol. 21, pp. 592 - 606, 2002.

[5] A. R. B. Blog. October, 2010, The Dangers of SPAM. Retrieved: June, 2011. Available: http://www.anthonyricigliano.info/the-dangers-of-spam/

[6] A. C. Solutions. January 7, 2011 Statistics and Facts About Spam. Retrieved: July, 2011. Available: http://www.acsl.ca/2011/01/07/statistics-and-facts-about-spam/

[7] H. R. Cournane A, "An analysis of the tools used for the generation and prevention of spam," Computers & Security, vol. 23, pp. 154-66, 2004.

[8] R. JENNINGS. JANUARY 28, 2009, Cost of Spam is Flattening - Our 2009 Predictions. Retrieved: July, 2011. Available at: http://ferris.com/2009/01/28/cost-of-spam-is-flattening-our-2009-predictions/

[9] Y. Hu, et al., "A scalable intelligent non-content-based spam-filtering framework.," Expert Syst. Appl., vol. 37, pp. 8557-8565, 2010.

[10] J.-J. Sheu, "An Efficient Two-phase Spam Filtering Method Based on E-mails Categorization " International Journal of Network Security, vol. 9, pp. 34-43, July 2009.

[11] C.-H. Wu, "Behavior-based spam detection using a hybrid method of rule-based techniques and neural networks," Expert Systems with Applications, vol. 36, pp. 4321-4330, April, 2009

[12] M. Ye, et al., "A Spam Discrimination Based on Mail Header Feature and SVM," In Proc. Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on Dalian Oct. 2008.

[13] C.-C. Wang and S.-Y. Chena, "Using header session messages to anti-spamming," Computers & Security, vol. 26, pp. 381-390, January 2007.

[14] Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. "The WEKA Data Mining Software: An Update. SIGKDD Explorations", 2009.

[15] P. R. Network Working Group, Editor. "Request for Comments RFC 2822," Retrieved: July, 2011. Available: http://tools.ietf.org/html/rfc2822.html

[16] CEAS 2008 Live Spam Challenge Laboratory corpus. Retrieved: March, 2011. Available at: http://plg1.uwaterloo.ca/cgi-bin/cgiwrap/gvcormac/fooceas.

[17] R. Beverly and K. Sollins, "Exploiting Transport-Level Characteristics of Spam," presented at the CEAS, Mountain View, CA, August 2008.

[18] C. GROUP. (2010, Spam email datasets, CSDMC2010 SPAM corpus. Retrieved: March, 2011. Available at: http://csmining.org/index.php/spam-email-datasets-.html

[19] T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters - Special issue: ROC analysis in pattern recognition, vol. 27, pp. 861-874, June 2006

 ISBN: 978-1-61208-176-2

Fault Tolerant Distributed Embedded Architecture and Verification

Chandrasekaran Subramaniam
Research and Development
Rajalakshmi Engineering College/ AUT
Chennai, India
chandrasekaran_s@msn .com

Prasanna Vetrivel, Srinath Badri
Electrical and Electronics Engineering
Easwari Engineering College/ AUT
Chennai, India
prasannavetrivel1990@gmail.com,
srinath.badri@live.com

Sriram Badri
Electronics and Communication Engineering
Sri Venkateswara College of Engineering/ AUT
Chennai, India
b-sriram@hotmail.com

***Abstract*— The objective of the work is to propose a distributed embedded architecture model for tolerating faults while performing security functions using multiple field programmable gate arrays (FPGA). The hardware encryption and decryption modules are used as customized modules within the devices to act as a cooperative system to tolerate omission and commission faults. The different security functions are communicating through a common UART channel and security operations are synchronized with standard protocols initiated by an embedded micro controller. The decision in locating the working and available modules among a pool of devices is carried out by the microcontroller using an intelligent F-map mechanism and directs the control instructions. The model is scalable with increased number of similar devices when connected across the common communication channel. The model is verified for all its paths using Symbolic Model Verifier, NuSMV to assert the dynamic behavior of the architecture in case of different faulty conditions.**

Keywords- Distributed architecture; Fault tolerance; Security module; Model verification; Assertion technique.

I. INTRODUCTION

The security architecture of embedded systems depends not only on the functional and performance requirements but also on the cost and spatial requirements suited to the target platforms. For example, the data path should be secured against many privacy attacks in the case of embedded systems used in the mobile applications. The data flow based on the dependence graph is solely determined by the components embedded as intellectual properties to perform the expected computations in real time. The architecture suitable for such computations with the security primitive components should be formally verified in order to avoid security errors like communication and synchronization errors. Due to the heterogeneity of various security hardware components from different component vendors, integration of them may lead to further challenges in the architectural design. The earlier SAFES architecture [1] focuses on the reconfigurable hardware that monitors the system behavior to realize the intrusion detections. The other proposed SANES architecture [2] focuses on security controller and component controller components to monitor the abnormal behavior in the system run time. Irrespective of the cryptographic algorithm used in the security primitive, the primitive components are to be self tested since the data path may vary dynamically in the case of a distributed embedded system. The components may not be available at some point of time when they are in need and they should be available in a fault free condition within the distributed mesh of devices. The correct selection of the primitive components either in the source or in the destination FPGA is to be decided in a power efficient manner among a pool of similar devices. The security processes are to be completed in the correct sequence and the operations are to be enabled as in the same dataflow form when they are completed [3]. The components are treated as resources within a single device to complete the submitted task and other similar devices are considered as coordinating devices controlled by a centralized controller. A field mapping technique is proposed through which the resource components available in the devices will be connected to form a distributed system considering the power consumption and the propagation delay involved in the on demand architecture. The distributed security architecture model has to be formally verified for its behavior to meet the reachability and fail free conditions. The standard NuSMV tool [5] supports LTL model checking [5] where the individual parameters can be inspected to investigate the effect of choices. Existing embedded system architectures are not capable of keeping up with the computational demands of security processing, due to increasing data rates and complexity of security protocols [7]. High assurance cryptographic applications require a design to be portioned and physically independent to ensure information cannot leak between secure design partitions. Ensuring this partition separation in the event of

 ISBN: 978-1-61208-176-2

independent hardware faults is one of the principle tenets of a high-assurance design [8]. FPGAs are highly promising devices for implementing private-key cryptographic algorithms. Compared to software based solutions FPGA [9] based implementations can achieve superior performance and security. Hence the main focus of the work is to propose a fault tolerant distributed architecture model for security hardware and check the model for its expected behavior meeting the specifications or against it.

The organization of the paper is as follows: Section 2 proposes the distributed embedded architecture for the security hardware using a single reconfigurable FPGA device with all the needed security primitive components. Section 3 discusses the sequences of processes to meet the performance requirements of the security architecture. The next section 4 extends the single system on chip to multi FPGA model where the data packet is routed between four similar devices under different faulty situations that lead to the worst case performance of the model. The next section explains the role of the micro controller in managing the field mapping of the devices when needed security components are faulty. Section 7 illustrates the model checking using NuSMV and its verified output and explores the scalability of the architecture along with its limitations like the issue of packet conflicts along the communication path in UART.

II. DISTRIBUTED EMBEDDED SYSTEM ARCHITECTURE

The architecture selection depends on the resource availability and reliability requirements of the security system. A reconfigurable platform like FPGA in addition to a high speed micro controller can organize the different computations needed to control the sequence of operations in the system. The resources are available in the form of intellectual properties (IPs) within the chip and these are to be assigned with the tasks in an efficient manner by the micro controller as the central manager. The best suitable architecture is the distributed embedded architecture where the resources are different sets of workers and the operations are performed in data flow form. The basic security hardware architecture consists of an encryption and decryption modules for 64 bit size along with their corresponding activation switches. The data transmission and reception takes place through the transmitter and receiver modules placed in the same chip and configured for different baud rates and actual process takes places through a universal asynchronous receiver transmitter (UART) module. The configuration of the security modules are self tested by the wired built-in self test (BIST) components connected with them to tolerate functional faults and controlled by a BIST Controller to regulate the data flow. All the components are selected or enabled by the control signals from device selector (DEVICE SEL) as shown in Fig. 1. The device reads the plain text from the keyboard and routes through the encryption and decryption modules as per the instructions received the micro controller that may reside inside or outside the chip assembly and connected through a bus. The deciphered text will be displayed in the activated display unit of the selected chip.

The micro controller is responsible for maintaining the queue of tasks and allocates the tasks to computational module as and when they are fail free. Because of the different execution times needed for different resources for processing the tasks, the micro controller has to run an intelligence algorithm to decide the available resource for that instant of time to forward the data. The micro controller gets the updated components information by regular sample intervals and refreshes its resource status (RST) and device status table (DST) in the private registers. The control signals are issued by the device select component as and when needed to send or receive. In case where there are concurrent requests from multiple devices over the limited bandwidth channel, then the micro controller forms a priority table to decide the order in which the queue of tasks may be completed. The micro controller is embedded with the algorithm to assign the priority to various requests.

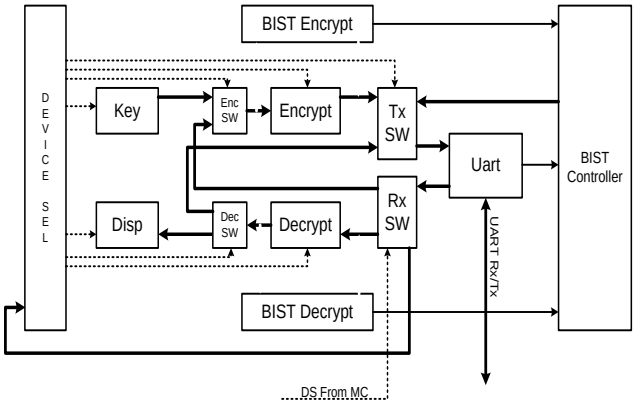


Figure 1. Distributed Embedded Architecture on Chip.

III. SEQUENCE OF PROCESSES

The security components needed to perform cryptographic operations are to be initialized with requests from the central manager component that is the micro controller. The sequence of processes is determined based on the availability of those components in their fault free conditions so that the data flow can be triggered. The enabling and disabling of the components in the pool of resources is taken care by the intelligent algorithm called F-Map embedded in the micro controller. For encryption of the text that is coming through an input peripheral say, keyboard connected to any one of the FPGA device, the sequence of processes is different from that of the decryption processes towards the output peripheral say, display device connected to other or the same FPGA device. For any FPGA to work it is mandatory that the Transmission, Reception switches & the UART connection to function properly. The microcontroller maintains two

 ISBN: 978-1-61208-176-2

table a Device Status Table (DST), where a Farm of FPGAs that can take part in the security process are listed and Resource Status Table (RST), where the resource status of each FPGA in the device farm is listed. During runtime micro controller forms a Run Time Table in which all the working set of FPGAs, obtained from DST with all its resources in fault-free condition (resource status obtained from the RST) for the current operation i.e. encryption or decryption, are enlisted depending on power dissipation (Power Aware mode) or propagation delay (Performance mode) of the FPGAs.

A. *Sequence of Operation*

- Check for updates in the run-time table_encryption
- Get data from the SOURCE_FPGA.
- Check the status of resources in the SOURCE_FPGA.
- If Fault free, send data to FPGA for encryption.
- Else, select next FPGA based upon the mode of operation in the runtime table to encrypt the data.
- Check updates in run-time table_decryption.
- Check the status of resources in the DESTINATION_FPGA
- If Fault free, send encrypted data to DESTINATION_FPGA for decryption.
- Else, select next FPGA in the runtime table and instructs to decrypt
- Transmit data to destination.
- Wait and Go to 1.

IV. MULTI FPGA BASED FAULT TOLERANT DISTRIBUTED EMBEDDED SYSTEM

The distributed embedded architecture on chip proposed in the work can be integrated with similar devices so as to make a distributed fault tolerant model. A micro controller is connected to instruct and manage the resource allocation between the devices based on the fault free conditions of the needed components. The algorithm that is embedded in the controller accepts the status of all the devices in terms of their resource components and decides the routing that the plain or cipher text has to follow as shown in Fig 2. The reliability of the distributed system is enhanced by component dynamic redundancy technique as and when the fault gets detected. Even though the functional component is static within a device, the controller calls the fault free components dynamically. The availability and system reliability is enhanced at the cost of communication overhead between the controller and the status table register multiple times. Since the performance depends on the speed of completion of the submitted task, the reliability of the system gets improved over the expected performance level.

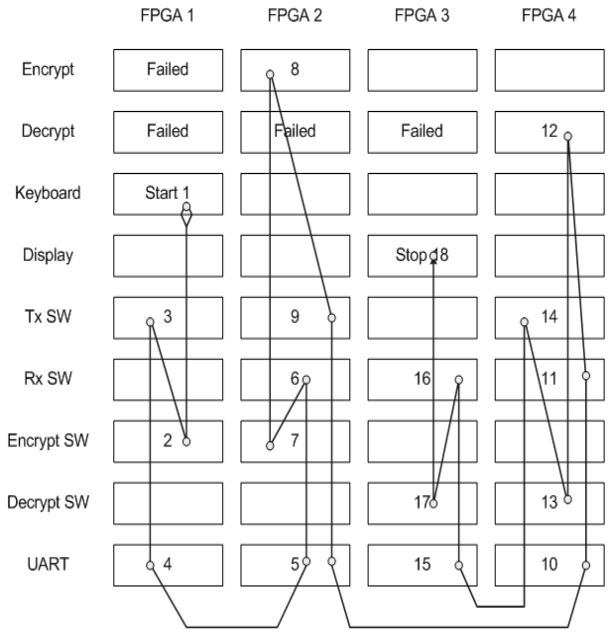


Figure 2. Multi FPGA based Distributed Embedded System.

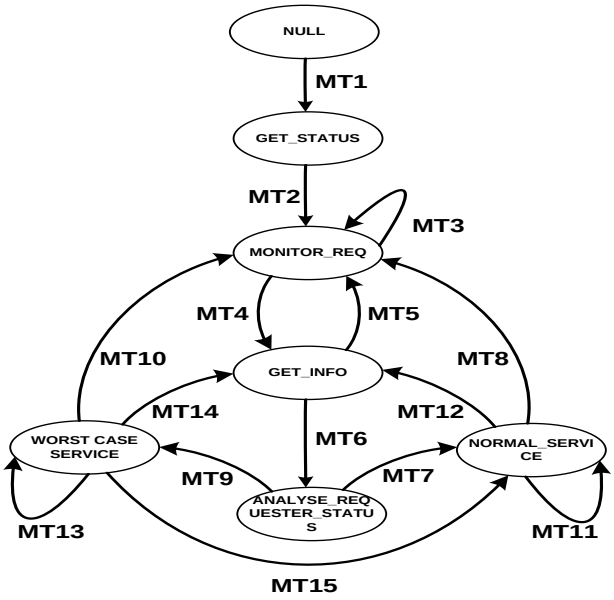


Figure 3. Microcontroller System Level Behaviour.

In the distributed architecture, the microcontroller behavior starts from instructing the distributed architectutre of FPGAs in its Null state. After a time interval (when an implicit or watchdog timer expires) its state transits from Null to Get_Status state (MT1). The state changes from Get_Status to Monitor_Req when the microcontroller successfully reads the status of all the FPGAs (MT2). The coordinating micro controller continues to remain in

Monitor_Req state until any request is made (MT3). The state transits from Monitor_Req to Get_Info if a request is made by any one of the FPGAs (MT4). On occurance of any data error during the request, the controller tracks back from Get_Info to Monitor_Req (MT5). On successfully aquiring the information from FPGAs, a state transition occurs from Get_Info to Analyse_Requester_Status (MT6). If all the resource primitive components at the source and destination FPGAs are fault-free, a state changes from Analyse_Requester_Status to Normal_Service (MT7). On successful completion of the task state transits from Normal_Service state to Monitor_Req state (MT8). If source or destination devices or both have non availability errors in any of their resources, the controller changes state from Analyse_Requester_Status to Worst_Case_Service (MT9). On successful completion of the event, the state changes from Worst_Case_Service state to Monitor_Req (MT10). The control remains in the Normal_Service state until the predefined time elapses on occurance of a communication error (MT11). The microcontroller changes state from Normal_Service to Get_Info state if the error continues to persist at the end of the predefined time interval (MT12). While encountering various errors, the control remains in the Worst_Case_Service state until the predefined time elapses (MT13). The control changes Worst_Case_Service to Get_Info state if the error exists when the predefined time is exhausted (MT14). The control of operations shift from Worst_Case_Service to Normal_Service when the micro controller finds the updated status of the resouce primitive components at the source and destination FPGAs to be fault-free (MT15) as in Fig 3. The fault tolerant technique using available fault free components can be extended to the situation when multiple FPGA devices are interconnected. Assuming similar devices, the security operations are executed in different devices based on the primitive components availability. If not, the correct routing instruction will be issued by the micro controller based on its updated device table and corresponding resource tables. The reliability of the communication path between the devices is a major challenge in the design of the above multi FPGA based distributed embedded security system.The synchronization in the execution of primitive operations is taken care by the embedded algorithm residing in the micro controller.

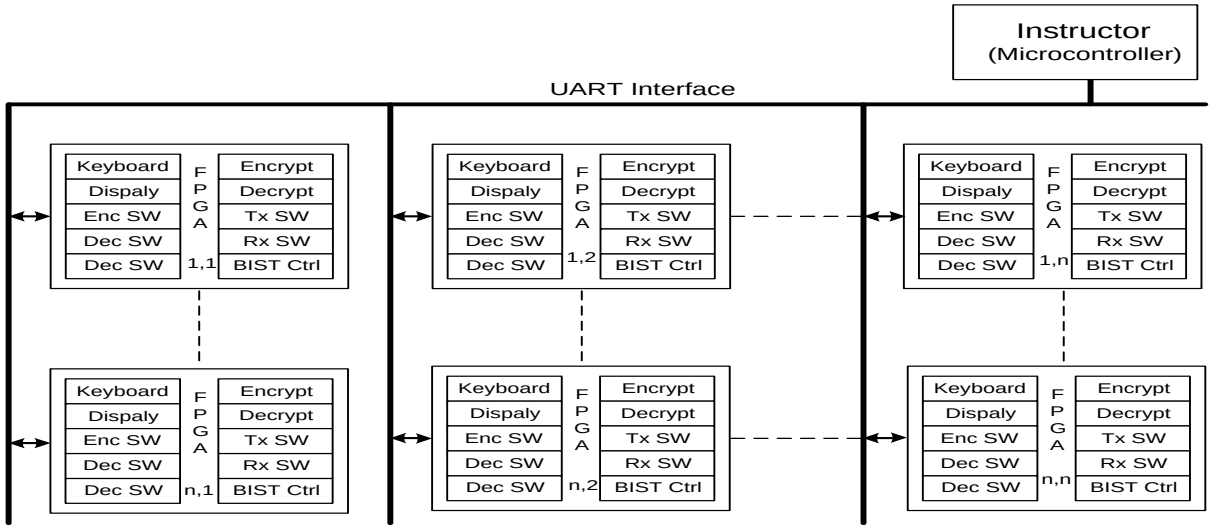


Figure 4. F Map for Power Aware Application.

V. F MAP FOR RESOURCE SELECTION

The Farm map (F-Map) is an intelligent algorithm by which the micro controller identifies the next device that has the needed security primitive components. The mapping brings the current status of the resources and fills the cell as '1' in the *START FPGA* which indicates that the *DEVICE SELECT* and the *BIST for Encryption* components are in failed states as shown in Fig 5. The algorithm searches the next device, depending upon whether the FPGAs are operated in power save mode or performance mode and in which all the needed components are fault free that is indicated by 1 cell for encryption process. Similarly the same algorithm is iterated for the decryption process also to get the data packet encrypted and communicated to the destination. The decryption may take place in the destination FPGA device. If multiple '1' cells are appearing at any time among different devices, then the controller selects the shortest path to minimize the communication delay to avoid any attack on the device itself that is possible in the distributed FPGA architecture. The *Device Select* at the destination FPGA is in failed state which makes the micro controller iterate for the next FPGA with the required resources in Fault-free state. The alternate FPGA for decryption process is selected depending on whether the distributed FPGA architecture is operated in performance mode or power aware mode. ACTIVE_DECRYPT_FPGA1 is chosen when operating in performance aware mode and

 ISBN: 978-1-61208-176-2

ACTIVE_DECRYPT_FPGA2 is chosen while operating in power aware mode. The performance characteristics of the distributed FPGA vary with the operating mode. For quicker response, the Distributed architecture may be operated in performance mode which causes higher power dissipation in the device. While operating in power save mode, the propagation delay is high, but the power consumed is minimal, as shown in Table. 1.

Propagation Delay = n * T (1)

Where, n= R + C - 1

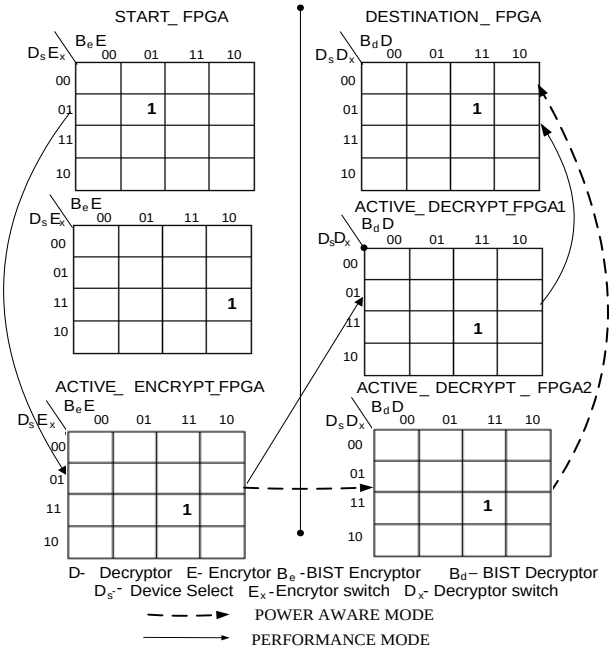


Figure 5. F Map for Power Aware Application.

VI. PERFORMANCE IN BEST CASE AND WORST CASE SCENARIO

In the distributed embedded system, if all the resources in the source and destination FPGAs are fault-free then the micro controller chooses the normal operating mode which is the best case scenario. The behavior of the embedded system during its best case performance is shown in Fig. 6 and its transitions are mentioned in Table. 2. The behavior of the distributed embedded system varies when a resource in source or destination FPGA is in failed state, depending upon the mode of operation viz. Performance Mode, Power Aware Mode, Safe Mode, depicted by the worst case behavior is shown in Fig. 7.

The microcontroller iterates various paths from the source to the destination depending upon the mode of operation and selects the FPGAs in source or destination side, with its required resources in fault free condition.

'R' Denotes the Row & 'C' Denotes Column in which the FPGA is located in the DST.

TABLE I. POWER DISSIPATION AND PROPOGATION DELAY FOR VARIOUS MODES

Order in DST (R,C)	Power Dissipation	Propagation Delay(ms)	Remarks
1,1	6.41W	T	Performance
2,3	4.23W	4*T	Normal/Safe
1,9	3.47W	9*T	Power Aware
6,12	7.9W	17*T	Poor

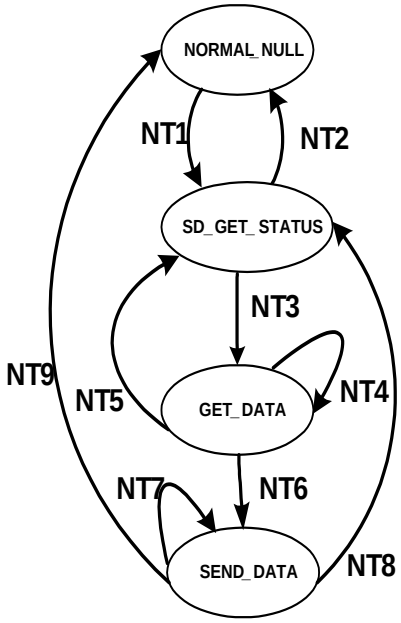


Figure 6. Best Case Scenario.

TABLE II. BEST CASE TRANSITIONS

Transition ID	Events
NT1	If main state machine reaches NORMAL_SERVICE state
NT2	Failure of getting source status || Failure of getting destination status
NT3	After getting successful status of source and destination
NT4	If not getting correct data from source
NT5	If not getting correct data from source && reaches maximum tries
NT6	Getting successful data from source
NT7	Not sending correct data to destination
NT8	If not sending correct data to destination
NT9	If sending correct data to destination

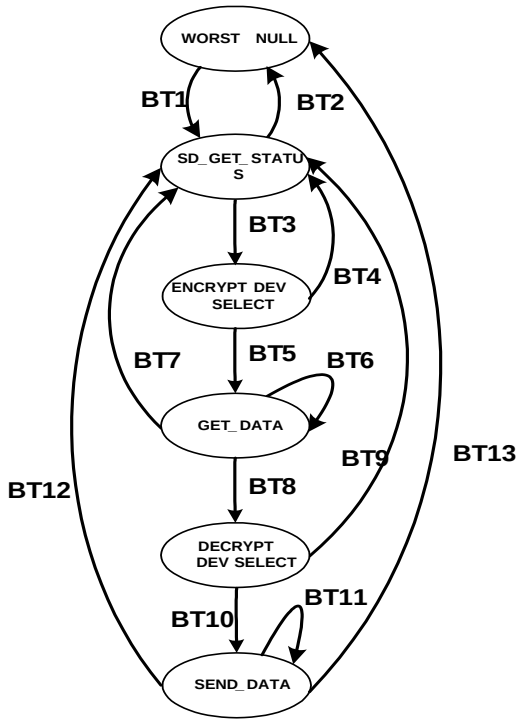


Figure 7. Worst Case Scenario.

While the distributed Embedded system is in its Worst Case behavior, the microcontroller acquires the status of all the FPGAs present in the distributed embedded system (BT1). In case of failure of acquisition of the status, the microcontroller waits in a null state until the status of the FPGAs are updated (BT2). Depending upon the mode of operation during resource failure in the source FPGA, the microcontroller searches for an alternate FPGA for encryption process(BT3). On selection of a suitable FPGA, the data from the source FPGA is transferred for encryption process (BT5). The microcontroller continues to exist in Get_Data state until the data is completely transferred to the alternate FPGA (BT6). On occurance of any errors during the data acquisition process, the microcontroller tracks back to its initial state (BT7). On completion of encryption process, a FPGA is selected by the microcontroller for decryption on the destination side (BT8). On successful selction of FPGA with all its resources in fault free condition, process of decryption takes place (BT10). The microcontroller traces itself to its initial state on failure of the decryption process (BT9). The decrypted data is ultimately sent to the destination FPGA until which the microcontroller remains in its current state (BT11). On successful reception of the complete decrypted data by the destination FPGA, the microcontroller prepares itself for the next encryption process (BT13), else the microcontroller resets to repeat the failed process (BT12).

VII. MODEL CHECKING OF ARCHITECTURE USING NUSMV

There are two extreme cases in which the fault tolerant capability can be measured. In one case, all the needed security components are fail free and available in a single chip and thereby the communication delay will be only due to transmit and receive signal propagation making power consumption very high. The other extreme is when one set of encryption components may be available in nearby device where as the decryption set of components in other end of the array or pool of devices and vice versa [6]. This condition leads to huge delay due to multiple enable, multiple address and data signals along with multiple receive, transmit and multiple device select signals. The worst case performance where the micro controller searches for the fail free encryption and decryption components is verified using the symbolic model verifier as shown below. The model checking is done to check the reliability of the model during the worst case performance and to verify the availability of alternate resources to complete the specified task. The LTL and CTL operations are applied as specifications to verify the availability of devices and the needed security resources for the encryption and decryption processes as given in NuSMV code below:

NUSMV CODE FOR AVAILABILITY IN THE WORST CASE SCENARIO

```
MODULE main
VAR
state : { null_state, get_update_status, s_device_select,
get_data, d_device_select, send_data };
status: boolean;
source_device_select : boolean;
data_get : boolean;
data_sent : boolean;
time : boolean;
dest_device_select : boolean;
INIT
state=null_state;
ASSIGN
next(state):= case
state = null_state : get_update_status ;
state = get_update_status & !status : null_state;
state = get_update_status & status : s_device_select;
state = s_device_select & !source_device_select :
get_update_status; state = s_device_select &
source_device_select : get_data;
state = get_data & !data_get & time : get_data;
state = get_data & !data_get & !time : get_update_status;
state = get_data & data_get : d_device_select;
state = d_device_select & !dest_device_select :
get_update_status;
state = d_device_select & dest_device_select : send_data;
state = send_data & !data_sent & time : send_data;
```

```
state = send_data & !data_sent & !time : get_update_status;
state = send_data & data_sent : null_state;
esac;
SPEC
EF (state=null_state)
SPEC
AG AF( !time -> state= get_update_status)
SPEC
AG EF(!status -> state=null_state)
SPEC
AG AG AF( time -> !data_get ->state= get_update_status)
```

The model checking algorithm reports true if specification holds true for every state of the system model. Otherwise, states not satisfying the specification are identified. A transition path from a defined initial state to a state identified as not satisfying the specification is called counterexample [4]. The CTL properties of the proposed security farm architecture is verified using NuSMV for different specifications and the result with instances and counter example is shown below:

SPECIFICATUIONS FOR THE NUSMV CODE USING LTL AND CTL PROPERTIES

```
-- specification EF state = null_state is true
-- specification AG (AF (!time -> state = get_update_status)) is true
-- specification AG (EF (!status -> state=null_state)) is true
-- specification AG (AG (AF (time -> (!data_get -> state = get_update_status)))) is false
-- as demonstrated by the following sequence
Trace DEscription: CTL Conterexample
Trace Type: Counterexample
-> State: 1.1 <-
 state = null_state
 status = FALSE
 source_device_select = FALSE
 data_ get = FALSE
data_sent = FALSE
 time = FALSE
 dest_device_select + FALSE
-> State: 1.2 <-
 state = get_update_status
 status = TRUE
-> State: 1.3 <-
 state = s_device_select
 status = FALSE
 source_device_select = FALSE
 time = TRUE
-- Loop starts here
-> State: 1.4 <-
 status=get_data
 source_device_select = FALSE
-> State: 1.5 <-
```

VIII. CONCLUSION AND FURUTE WORKS

A distributed embedded architecture model for security hardware with fault tolerant features is proposed to tolerate non availability faults and resource component faults. The model is a centralized system in which sequential control is exercised by a micro controller through the F-map algorithm embedded in it and can be scaled as multi FPGA devices pool with power or performance awareness. The collective behavior of the architecture model is formally verified using a model checker based on LTL and CTL properties. The limitation of the proposed model is in the communication overhead when all devices want to send the data concurrently over the limited bandwidth of UART channel. The constraints are the assumptions that the basic transmission and reception switches must always be fault free even to intimate the error report to the central controller. The actual implementation of the project with multiple FPGAs and triple micro controllers to enhance the reliability is the future work planned.

REFERENCES

[1] G. Gogniat, T. Wolf and W. Burleson: *Reconfigurable Hardware for High-Security / High-Performance Embedded Systems: The SAFES Perspective*, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 16, No. 2, February 2008, pp. 1-10.

[2] *Reconfigurable Security Architecture for Embedded Systems*, http://vcsgi.ecs.umass.edu/essg/papers/MOCHASubmit.pdf, pp. 1-7.

[3] S. Hauck and A. Dehon: Morgan Kaufmann Publications, *Reconfigurable Computing*, pp 107-110.

[4] Sebastian Steinhorst: Dissertation, *Formal Verification Methodologies for Nonlinear Analog Circuits*, Frankfurt, 2011, pp. 55-69.

[5] A.Cimatti, E.Clarke, F.Giunchiglia and M.Roveri: *NuSMV :A New Symbolic Model Verifier*,pp.1-5.

[6] G. K. Palshikar: *An Introduction to Model Checking*, html page, pp. 1-8.

[7] S. Ravi, A. Raghunathan, P. Kocher and S. Hattangady: *Security in Embedded Systems: Design Challenges*, ACM Transactions on Embedded Computing Systems, Vol 3, No. 3, August 2004, pp. 6-10.

[8] P. Quintana: *Fail-Safe FPGA Design Features for High-Reliability Systems*, Paper ID: 900566, IEEE 2009, pp. 3-5.

[9] A. Dandalis, V.K. Prasanna and D.P. Rolim: An Adaptive Cryptography Engine for IPSec Architectures, ACM Transactions on Design of Automation of Electronic Systems, Vol. 9, July 2004, pp. 333-353.

 ISBN: 978-1-61208-176-2

Determining Authentication Strength for Smart Card-based Authentication Use Cases

Ramaswamy Chandramouli
Computer Security Division, Information Technology Lab
National Institute of Standards and Technology
Gaithersburg, MD, USA
mouli@nist.gov

Abstract - Smart cards are now being extensively deployed for identity verification(smart identity tokens) for controlling access to Information Technology (IT) resources as well as physical resources. Depending upon the sensitivity of the resources and the risk of wrong identification, different authentication use cases are being deployed. Assignment of authentication strength for each of the use cases is often based on: (a) the total number of three common orthogonal authentication factors – What You Know, What You Have and What You are – used in the particular use case and (b) the entropy associated with each factor chosen. The objective of this paper is to analyze the limitation of this approach and present a new methodology for assigning authentication strengths based on the strength of pair wise bindings between the five entities involved in smart card based authentications – the card (token), the token secret, the card holder, the card issuer and the person identifier stored in the card The use of the methodology for developing an authentication assurance level taxonomy for a real world smart identity token deployment is also illustrated.

Keywords - Identity Verification; Smart Identity Token; Authentication Strength

I. INTRODUCTION

Smart cards are now being extensively deployed for identity verification for controlling access to Information Technology (IT) resources as well as physical resources [1,2,3]. We refer to them as Smart Identity Tokens and use the two terms interchangeably throughout this paper. These types of smart cards generally carry: (a) A Person Identifier (PI), (b) A Secret (TS) usually in the form of a cryptographic key [4], and (c) A Credential linking the Secret and the Identifier (CR). Along with these data, a PIN (a combination of numbers) is often used for: (a) Activating the card (token) and for (b) Restricting access to certain data objects and operations. In some instances, presentation of a live biometric data (such as a fingerprint) is used to enable the above functions instead of a PIN. In any enterprise deploying smart cards, there may be different types of resources that may have to be protected by restricting access to only those whose identity is verified through a smart card based authentication mechanism. Depending upon the sensitivity of the resource and the risk associated with wrong identification of the entity requesting access to those resources, authentication mechanisms using different combinations of the three data types enumerated above (PI, TS or CR) along with/without an activation data may be used. A set of authentication mechanisms used by an enterprise for controlling access to different types of resources (or stated differently- different applications of smart identity token) are called Authentication Use Cases.

In general (irrespective of whether a smart identity token is used or not), the choice of an Authentication Use Case in the context of an access control to a resource is often made based on authentication strength or assurance level associated with the token artifact used in the Authentication Use Case. These artifacts are: (a) an identifier specific to a domain and (b) a credential that is a combination of an identifier and a secret – examples for the latter being: (a) a PIN (b) a one-time password and (c) a cryptographic key. The usage of a token by a claimant during an authentication event results in a value called Authenticator that is generated by the token and is transmitted from the token to the authentication module or the verifier. The basis for designating an authentication strength associated with a token is a fundamental unit called "Authentication Factor". There are three main authentication factors [5]:

- What the Entity Knows (e.g., Password, PIN, etc)
- What the Entity Has (e.g., possession of a token that generates one-time passwords)
- What the Entity Is (e.g., inherent physiological characteristic such as a Fingerprint)

A token that uses one of the above three factors is called a single factor token (e.g., a password that belongs to " What the Entity Knows" factor). A token that uses a combination of two or more of the above factors is called a multi-factor token. A smart card that contains an embedded private cryptographic key (thus using What the Entity Has authentication factor) that can be used to generate an authenticator when it is activated by a PIN (using the What the Entity Knows authentication factor) is deemed a multi-factor token. An authentication use case may use one or more tokens and hence may involve the use of one or more authentication factors. In general, the authentication strength associated with an authentication use case is determined based on the combination of the following metrics:

 ISBN: 978-1-61208-176-2

- The number of authentication factors used in the authentication use case
- The Entropy associated with each of the authenticator factor used

In this paper, we argue that the logic for assigning authentication strength based on the number of authentication factors in an authentication use case is valid only under certain limiting conditions and that these conditions do not hold in the case of authentication use cases using smart cards as identity tokens. This is the rationale for proposing a new methodology for assigning authentication strengths for various authentication use cases involving smart identity tokens.

The description of the conditions under which the number of authentication factors can be used as a reliable metric for authentication strength and an illustration of how those conditions do not hold in the case of smart cards are given in Section II. Section III discusses the basis vector that is applicable for smart card based identity verification approaches. The development of our methodology for determining authentication strengths for various smart card-based authentication use cases based on the basis vector referred to above is the topic of Section IV. The application of this methodology for assigning authentication strengths for building a taxonomy of authentication assurance levels for the set of authentication use cases specified for a major government smart card-based identity verification deployment is done in Section V. Section VI presents the benefits of our methodology and provides the conclusions.

II. LIMITATIONS OF AUTHENTICATION FACTOR-BASED APPROACH FOR DETERMINING AUTHENTICATION STRENGTHS

In order that the number of authentication factors is a valid metric for determining the authentication strength of an authentication use case, it must satisfy the following properties:

- AF-AS-P1: The authentication factors must be mutually independent. If there is any mutual dependency between any two authentication factors, then assuming the additive property is not valid for computing the metric indicating the authentication strength. This is not an issue as the three authentication factors – What You Know, What you Have and What you Are do not have any pair wise mutual dependency.
- AF-AS-P2: All authenticators used in the authentication use case must flow directly from the claimant to the verifier in the resulting authentication message protocol. This property must hold since any authentication decision by the verifier is based entirely on the outcome of the process of verifying one or more authenticators received from the claimant. Hence any authentication decision based on a lesser number of authenticators is certainly of lower authentication strength than an authentication decision using a higher number of authenticators.

We illustrate through an example that the second property is not satisfied in many smart card based authentication use cases deployed in real-world implementations [3,8]. For example, in an authentication use case called Challenge-Response, the smart card responds to a random challenge string sent by the authentication system by encrypting the string with its private key and sending the encrypted string back. Some cards are programmed to require the card holder to provide a PIN to perform this private key operation. This authentication use case is classified as a two factor authentication (since it involves demonstrating the presence of a secret cryptographic key (one factor) and the PIN (second factor)) although the only authenticator that flows to the authentication system (verifier) is the encrypted challenge. Thus, we see that, in order to truly assess the authentication strength associated with smart card based authentication use cases, we need a basis vector other than just the number of authentication factors. To identify and derive such a basis vector, we find that there is a need to look at the various basic entities that participate in authentication protocols using smart cards and the nature of pair-wise binding that exists among them. The logic for development of these pair wise bindings is described in the next section.

III. DEVELOPMENT OF BASIS VECTOR FOR SMART CARD-BASED AUTHENTICATION USE CASES

Before we start using the pair-wise binding as components of a basis vector used for determining authentication strengths, we need to make a comprehensive list of the basic entities involved in them. These basic entities, building on the smart card contents we saw in the last section are: the physical token (smart card), the card holder, the token secret, card issuer and the person identifier. Please note that we do not term the credential as a basic entity since credential is a derived artifact providing the binding of the two basic entities – Person Identifier and the Token Secret. Before we start listing the pair-wise bindings, we find that any authentication use case is itself built from some primitive authentication usage modes each of which uses one or more of three categories of smart card data – Person Identifier, Token Secret and Credential. Hence every pair-wise binding should trace its link to a primitive authentication usage mode and the associated smart card data used in that mode. This link is provided through the data in Table I. Table I, in addition to providing the bindings, also provides the strength associated with each binding based on the nature of the primitive authentication usage mode and the associated data used in it. Out of the six possible valid bindings, the person identifier participates in three of them being associated with card issuer (through digital signature), token secret (being used in digital certificate) and card holder (being used in biometric object).

 ISBN: 978-1-61208-176-2

TABLE I. SMART IDENTITY CARD – PRIMITIVE AUTHENTICATION USAGE MODES & BINDINGS

Smart Card Data	Primitive Authentication Usage Mode	Pair-wise Bindings with associated strength
Embedded Cryptographic Key (private key of an asymmetric Key Pair) – Token Secret	PUM-1: Verifying Presence of embedded token secret (tested by sending an input data from the Verifier and receiving an associated Authenticator)	Token- Token Secret Binding (Strong)
Embedded Cryptographic Key (private key of an asymmetric Key Pair) – Token Secret that requires an activation data to demonstrate its presence	PUM-2: Same as previous + card holder providing a PIN for generating the authenticator	1.Token – Token Secret Binding (Strong) 2.Card Holder – Token Binding (Strong or Weak depending upon entropy of activation data)
Person Identifier	PUM-3: Person Identifier's origin and integrity checked using its associated digital signature	Person Identifier- Card Issuer Binding (Strong)
Credential (A Public Key Certificate) linking the token secret to the Person Identifier	PUM-4: Trust in the certificate established through Certificate Validation	Token Secret – Person Identifier Binding (Strong)
Credential (A digitally signed Biometric Object) linking a Card Holder Trait (biometric) to the Person Identifier	PUM-5: The digital signature associated with biometric data object is verified. Live biometric sample sent to the card for matching with the stored biometric data	Card Holder – Person Identifier Binding (Strong or Weak depending upon how live sample is collected)

IV. METHODOLOGY FOR ASSIGNING AUTHENTICATION STRENGTHS FOR AUTHENTICATION USE CASES

In the previous section, we identified the primitive authentication usage modes and the bindings (along with their associated strength) enabled by those modes. An authentication use case that is used in a smart identity token deployment will be a combination of one or more of the primitive authentication usage modes. Now our final goal is the determination of authentication strength for a given authentication use case. In order to compute this value, we need to know the security properties satisfied and the weakness in each of the primitive authentication usage modes that constitute that authentication use case. The derivation of these security properties satisfied and weaknesses from the bindings (and their associated strengths) provided by each of the five primitive authentication usage modes (taking into consideration the state of smart card technology) is shown in Table II.

Now, based on the observation that the primitive authentication modes are independent of each other (except for PUM-2 which is a superset of PUM-1), the security properties associated with the set of primitive authentication usage modes constituting an authentication use case can all be added up to obtain the total set of security properties satisfied in an authentication use case.

Let us consider the following Authentication use case which we shall call as BIO-A:

1. The Authentication Module (Verifier) reads the signed biometric object on the card.
2. The digital signature of the biometric object is verified.
3. The Authentication station is attended by a guard under whose watch the claimant provides his /her fingerprint through a scanner present in the station.
4. The Live sample of the biometric is compared with the stored biometric data on the card.
5. When the match is successful, the person identifier extracted from the signed biometric object is compared with identifier stored in the identifier object on the card. The digital signature associated with identifier object is verified.
6. If the verification is successful, the identifier is sent to the Physical Access Control Server which in turn sends a signal to open the door leading to the facility controlled by the authentication station.

From the description of the above steps in our example Authentication Use Case BIO-A, we find that steps 1-4 map to our primitive authentication usage mode PUM-5. Step 5 maps to our usage mode PUM-3. Hence adding the properties associated with these primitive authentication usage modes, we find that the authentication use case BIO-A satisfies the following total set of properties:

1. Card Holder is authenticated (Strong – since the live sample is collected under a supervised condition ensuring freshness and hence no replay using duplicated fingerprints possible)
2. Validity of the Identifier is established

The security property set associated with an authentication use case can be used as a metric for establishing a partial order among the various authentication use cases specified for a smart card based identity verification deployment scenario. This partial order can then be used to construct an authentication assurance level taxonomy for that deployment instance.

 ISBN: 978-1-61208-176-2

TABLE II. SECURITY PROPERTIES OF AUTHENTICATION USAGE MODES

Primitive Authentication Usage Mode	Bindings Established with associated strength	Security Properties Satisfied (WEAKNESS in CAPS)
PUM-1: Verifying Presence of embedded token secret (tested by sending an input data from the Verifier and receiving an associated Authenticator)	Token- Token Secret Binding (Strong)	1.Card is Authenticated 1.STOLEN CARD 2. CARD HOLDER IS NOT AUTHENTICATED 3. NO LINK FROM TOKEN SECRET TO PERSON IDENTIFIER
PUM-2: Same as previous + sending an activation data to the token	(a)Token – Token Secret Binding (Strong) (b)Card Holder – Token Binding (Strong or Weak depending upon activation data)	1. Card is Authenticated 2. Card Holder is Authenticated (Strength based on Activation Data) 1.NO LINK FROM TOKEN SECRET TO PERSON IDENTIFIER
PUM-3: Person Identifier's origin and integrity checked using its associated digital signature	Person Identifier – Card Issuer Binding (Strong)	1.Validity of the Identifier is established 1.STOLEN CARD 2. CLONED CARD 3. CARD IS NOT AUTHENTICATED 4. CARD HOLDER IS NOT AUTHENTICATED 5. NO LINK FROM TOKEN SECRET TO PERSON IDENTIFIER
PUM-4: Trust is established on a credential (a public key certificate) linking embedded token secret and the person identifier through certificate validation	Token Secret – Person Identifier Binding (Strong)	1.Link from Token Secret to Person Identifier established 1.STOLEN CARD 2. CLONED CARD 3. CARD IS NOT AUTHENTICATED 4. CARD HOLDER IS NOT AUTHENTICATED
PUM-5: Trust is established on a credential (signed biometric object containing the identifier in addition to biometric data) by verifying the digital signature. Live biometric sample sent to the card for matching with the stored biometric data	Card Holder – Person Identifier Binding (Strong or Weak depending upon how live sample is collected)	1. Card Holder is Authenticated (Strength based on how live biometric sample is collected) 1.CLONED CARD 2. CARD IS NOT AUTHENTICATED

V. ILLUSTRATION OF METHODOLOGY FOR A REAL WORLD SMART IDENTITY TOKEN DEPLOYMENT

In this section, we illustrate the application of our methodology for assignment of authentication strengths for authentication use cases used in a real world smart identity token deployment scenario. The first step of our methodology is to express each authentication use case specified for the deployment in terms of our primitive authentication usage mode. This will automatically provide us the total set of security properties associated with that authentication use case. We then use the property set containment to derive a partial order among the authentication use cases and to finally derive an authentication assurance level taxonomy for the entire smart identity token deployment. The real world smart identity token deployment we have chosen for our illustration is the Implementation of Personal Identity Verification (PIV) smart card for controlling physical access to federal facilities and logical access to U.S government IT systems [7,8]. For the sake of space and brevity, we do not describe each of the authentication use cases in the PIV deployment scenario. We also do not illustrate the process by which our primitive authentication usage modes can be composed to obtain a PIV authentication use case. These liberties have been taken since our final goal is just to illustrate the use of our methodology for developing an authentication assurance level taxonomy. Table III below provides a compilation of all the PIV Authentication uses [8], the list of primitive authentication usage modes that comprise each authentication use case and total set of security properties satisfied by each authentication use case specified in a PIV deployment instance.

 ISBN: 978-1-61208-176-2

TABLE III. PROPERTIES SATISFIED BY PIV AUTHENTICATION USE CASES

PIV Authentication Use Case	Set of Primitive Authentication Usage Modes involved	Properties Satisfied
Authentication using PIV CHUID (CHUID)	PUM-3: Identifier's origin and integrity checked using its associated digital signature	1.Validity of the Identifier is established
Unattended Authentication using PIV Biometric (BIO)	PUM-5: Trust is established on a credential (signed biometric object containing the identifier in addition to biometric data) by verifying the digital signature. Live biometric sample sent to the card for matching with the stored biometric data PUM-3: Identifier's origin and integrity checked using its associated digital signature	1. Card Holder is authenticated (Weak) 2.Validity of the Identifier is established
Attended Authentication using PIV Biometric (BIO-A)	PUM-5: Trust is established on a credential (signed biometric object containing the identifier in addition to biometric data) by verifying the digital signature. Live biometric sample sent to the card for matching with the stored biometric data PUM-3: Identifier's origin and integrity checked using its associated digital signature	1. Card Holder is authenticated (Strong) 2.Validity of the Identifier is established
Authentication using PIV Asymmetric Cryptography (PKI-AUTH)	PUM-4: Trust is established on a credential (a public key certificate) linking embedded token secret and the person identifier through certificate validation PUM-2: Verifying Presence of embedded token secret (tested by sending an input data from the Verifier and receiving an associated Authenticator) (derived from PUM-1) + sending a activation data of robust strength to the token PUM-3: Identifier's origin and integrity checked using its associated digital signature	1.Link from Token Secret to Identifier established 2. Card Holder is authenticated (Strong) 3. Card is Authenticated 4.Validity of the Identifier is established
Authentication using Card Authentication Certificate Credential (PKI-CAK)	PUM-4: Trust is established on a credential (a public key certificate) linking embedded token secret and the person identifier through certificate validation PUM-1: Verifying Presence of embedded token secret (tested by sending an input data from the Verifier and receiving an associated Authenticator) PUM-3: Identifier's origin and integrity checked using its associated digital signature	1.Link from Token Secret to Identifier established 2. Card is Authenticated 3.Validity of the Identifier is established

Based on the property containment relationship between the various PIV authentication use cases, we derive a partial order and use that partial order to develop a complete authentication assurance level taxonomy. The taxonomy thus derived in shown in Figure 1 below:

 ISBN: 978-1-61208-176-2

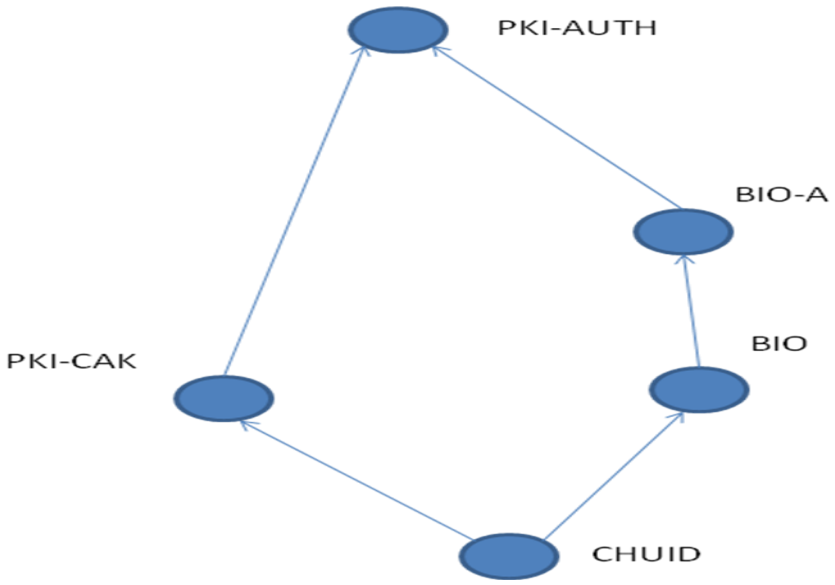


Figure 1. Authentication Assurance Level Taxonomy for a Smart Identity Token Deployment

VI. CONCLUSIONS AND BENEFITS

The observation that not all authenticators flow between the smart identity token and the authentication module (verifier) has driven the need for a new basis vector other than just the number of authentication factors and an associated methodology for assigning authentication strengths for various authentication use cases involving smart cards. In this paper, we developed such a methodology which uses pair wise bindings between the five entities involved in smart identity token authentication use cases –i.e., token (the card), the token secret, the card holder, the card issuer and the person identifier - as the basis for deriving a set of properties satisfied for each primitive authentication usage mode. The primitive authentication usage modes are in turn identified based on the types of data a smart identity token usually holds. Next, we illustrated the process of expressing an authentication use case in terms of the combination of primitive authentication usage modes and using the additive properties associated with each usage mode, derived the total set of properties satisfied by an authentication use case. Finally the property set associated with an authentication use case is used to derive a partial order among the use cases. This partial order was then used to derive an entire authentication assurance level taxonomy for a smart identity token deployment scenario. The advantages of this approach are: (a) It takes into account all entities participating in the authentication protocol (the five that we referred to earlier) and the pair wise bindings between them and (b) considers technology-specific weaknesses (e.g., token can be stolen and cloned) that may affect the security properties satisfied in each primitive authentication usage mode and by extension in an authentication use case.

REFERENCES

[1] Securing e-business applications using Smart Cards, IBM Systems Journal, Vol 40, Number 3, 2001, (Oct, 2011), http://www.research.ibm.com/journal/sj/403/hamann.html

[2] Kumar, M.: New Remote User Authentication Scheme Using Smart Cards, *IEEE Transactions on Consumer Electronics.* Volume 50, Issue 2, 597 – 600 (2004)

[3] TWIC Reader Hardware And Card Application Specification, May 30, 2008, (Nov, 2011) http://www.tsa.gov/assets/pdf/twic_reader_card_app_spec.pdf

[4] NIST SP 800-63-1 Recommendation for Electronic Authentication, Dec 2008, (Oct, 2011) http://csrc.nist.gov/publications/drafts/800-63-rev1/SP800-63-Rev1_Dec2008.pdf

[5] OMB M04-04 – E-Authentication Guidance for Federal Agencies, Dec 16, 2003, (Oct, 2011) http://www.whitehouse.gov/omb/memoranda/fu04/m04-04.pdf

[6] Internet X.509 PKI Certificate & CRL Profile, (Nov, 2011) http://www.ietf.org/rfc/rfc5280.txt

[7] Identity Management Task Force Report, National Science and Technology Council (NSTC) Subcommittee on Biometrics and Identity Management, 2008, (Oct, 2011) http://www.biometrics.gov/documents/idmreport_22sep08_final.pdf

[8] FIPS 201 – Personal Identity Verification of Federal Employees and Contractors**,** (Oct, 2011) http://csrc.nist.gov/publications/drafts/fips201-2/Draft_NIST-FIPS-201-2.pdf

 ISBN: 978-1-61208-176-2