IFIP AICT 399

Christos Douligeris
Nineta Polemi
Athanasios Karantjias
Winfried Lamersdorf
(Eds.)

Collaborative, Trusted and Privacy-Aware e/m-Services

12th IFIP WG 6.11 Conference
on e-Business, e-Services, and e-Society, I3E 2013
Athens, Greece, April 2013
Proceedings

Springer

IFIP Advances in Information and Communication Technology 399

IFIP – The International Federation for Information Processing

IFIP was founded in 1960 under the auspices of UNESCO, following the First World Computer Congress held in Paris the previous year. An umbrella organization for societies working in information processing, IFIP's aim is two-fold: to support information processing within its member countries and to encourage technology transfer to developing nations. As its mission statement clearly states,

> *IFIP's mission is to be the leading, truly international, apolitical organization which encourages and assists in the development, exploitation and application of information technology for the benefit of all people.*

IFIP is a non-profitmaking organization, run almost solely by 2500 volunteers. It operates through a number of technical committees, which organize events and publications. IFIP's events range from an international congress to local seminars, but the most important are:

- The IFIP World Computer Congress, held every second year;
- Open conferences;
- Working conferences.

The flagship event is the IFIP World Computer Congress, at which both invited and contributed papers are presented. Contributed papers are rigorously refereed and the rejection rate is high.

As with the Congress, participation in the open conferences is open to all and papers may be invited or submitted. Again, submitted papers are stringently refereed.

The working conferences are structured differently. They are usually run by a working group and attendance is small and by invitation only. Their purpose is to create an atmosphere conducive to innovation and development. Refereeing is also rigorous and papers are subjected to extensive group discussion.

Publications arising from IFIP events vary. The papers presented at the IFIP World Computer Congress and at open conferences are published as conference proceedings, while the results of the working conferences are often published as collections of selected and edited papers.

Any national society whose primary activity is about information processing may apply to become a full member of IFIP, although full membership is restricted to one society per country. Full members are entitled to vote at the annual General Assembly, National societies preferring a less committed involvement may apply for associate or corresponding membership. Associate members enjoy the same benefits as full members, but without voting rights. Corresponding members are not represented in IFIP bodies. Affiliated membership is open to non-national societies, and individual and honorary membership schemes are also offered.

Christos Douligeris Nineta Polemi
Athanasios Karantjias Winfried Lamersdorf (Eds.)

Collaborative, Trusted and Privacy-Aware e/m-Services

12th IFIP WG 6.11 Conference
on e-Business, e-Services, and e-Society, I3E 2013
Athens, Greece, April 25-26, 2013
Proceedings

Springer

Volume Editors

Christos Douligeris
Nineta Polemi
University of Piraeus
80 Karaoli and Dimitriou Str., 18534 Piraeus, Greece
E-mail: {cdoulig, dpolemi}@unipi.gr

Athanasios Karantjias
SingularLogic S.A.
Al. Panagouli and Siniosoglou, 14234 N. Ionia, Greece
E-mail: tkaratzias@singularlogic.eu

Winfried Lamersdorf
University of Hamburg
Vogt-Kölln-Str. 30, 22527 Hamburg, Germany
E-mail: lamersd@informatik.uni-hamburg.de

ISSN 1868-4238 e-ISSN 1868-422X
ISBN 978-3-642-37436-4 e-ISBN 978-3-642-37437-1
DOI 10.1007/978-3-642-37437-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2013934552

CR Subject Classification (1998): K.4, H.4, H.3, K.6.5, J.1, C.2, D.2

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume contains the revised papers of the 12th IFIP Conference on e-Business, e-Services, e-Society (I3E) which was held in Athens, Greece, April 25-26, 2013. The papers reflect the depth and breadth that the I3E conference series, now well in its second decade, attracts. Notwithstanding the inevitable problems and shortcomings of the ever-present e-society, it is undeniable that all of us, in one way or another, are directly affected, use, and enjoy it.

The international scientific community has been at the forefront of the developments in this area with several new conferences, journals, and books – in printed or electronic media – appearing every year. It is obvious from the papers in this volume that the I3E conference series has kept their leading position as one of the main venues for new ideas, practice presentations, and new research forum in the e-services academic and development society, having the full support and being the flagship conference of the IFIP WG6.11 "Communication Aspects of the E-World."

This year the focus of the conference was on "Collaborative, Trusted and Privacy-Aware e/m-Services" to indicate the fact that we live in world which is interconnected but requires a minimum level of security and privacy in order to fully exploit the benefits of e-society. This focus is apparent in the first two parts of the volume where trust and privacy issues are addressed in several e-services scenarios and with a variety of tools and technologies. In particular, the new paradigm of cloud computing raises a new slate of questions regarding security, privacy, and access control and poses a set of legal requirements that require immediate response. Social networking as a major form not only of communication and entertainment but of doing business on-line raises its own questions demanding new protocols and regulatory interventions. Unless we deal with adoption issues and go over a careful analysis of available and proposed e/m services, the scope of our work would be an abstract, theoretical model with no relevance to the real world. New ecological demands need to be addressed in using e-services and by using e-services. Business processes and knowledge management in the volatile e-society take a new meaning and require new approaches and innovative ideas. We need to create knowledge from data and raw information in an efficient, timely, and privacy-aware manner. Appropriate management policies are thus necessary to define how e-services can be adopted, expanded and continuously evaluated.

We would not be able to prepare this coherent, thorough, and up-to-date volume without the help of several individuals and organizations. We would like to extend our appreciation to our basic supporting organizations, IFIP and

Springer, for making these proceedings possible. The conference would not have been possible without the great financial and technical support of the University of Piraeus and the University of Piraeus Research Center. The Software and Integrated IT Solutions Group of Singular Logic SA provided much-needed financial and marketing during all the phases of this conference. The Greek Chapter of ISACA was instrumental in making the conference widely known in its community.

In addition, I3E 2013 can be considered the brainchild of the European project ImmigrationPolicy2.0, which has as its main objective to offer to legal immigrants, policy makers, public administrators, and experts in the domain of migration the capability to contribute their valuable knowledge and experience in order to effectively enhance current efforts of migration policy harmonization at a European level as well as to help identify existing gaps in national policies compared to E.U. directives. The project has been funded by the EU ICT PSP programme.

The conference included three keynote speeches by experts in the topics addressed by the conference. Andreas Mitrakas talked on "Information Technology Regulation: Towards a Future Proof Model?". Dimitris Karagiannis analyzed the topic "Open Models as a New Instrument in E-Business," while Günter Müller explored the question "Resilience: A New Paradigm for Security?". A roundtable on the uptake and acceptance of e-services run by George Doukidis was part of the conference program as well. Two workshops were also organized. The first, sponsored by the ImmigrationPolicy2.0 project and chaired by Thanos Karantjias, Stelios Pantelopoulos, and Dimitris Papanikas, explored e-migration issues and provided a live demonstration of the IMP2.0 platform. The second one was organised by Robert Woitsch and Wilfrid Utz and addressed the issue of Value Creation through Concept Models.

We would like to extend our warmest thanks to the Program Committee members who reviewed the papers, provided thorough reviews and incisive comments so that the camera-ready papers were of the highest possible quality. The authors of the papers, who come from several different nations and continents and have diverse academic and industrial backgrounds, need to be thanked for their trust in the conference and for their patience in following the guidelines provided by the reviewers and revise their papers accordingly. It should be mentioned that all the papers were evaluated by at least three independent reviewers according to several criteria and that the reviewers were asked to look at the papers critically regarding their presentation, content, and impact.

The Finance Chair and the local Organizing Committee are to be thanked for making all the necessary arrangements for a successful meeting in Athens.

We are certain that this volume will open new questions regarding e-services, their expansion, and the security and privacy concerns associated with their uptake as well as regarding the impact that these technologies may have on our everyday personal and business lives.

April 2013

Christos Douligeris
Nineta Polemi
Athanasios Karantjias
Winfried Lamersdorf

Organization

The 12th IFIP Conference on e-Business, e-Services, e-Society (I3E 2013) was organized by the University of Piraeus, Department of Informatics, and took place in Athens, Greece (April 25–26, 2013).

The conference was sponsored by IFIP WG 6.11 in cooperation with the University of Piraeus Research Center, SingularLogic S.A., ISACA and ImmigrationPolicy2.0 FP7 Programme.

General Chairs

Nineta Polemi	University of Piraeus, Greece
Georgios Doukidis	Athens University of Economics and Business, Greece

Steering Committee

Winfried Lamersdorf	University of Hamburg, Germany
Wojciech Cellary Poznan	University of Economics, Poland
Reima Suomi	University of Turku, Finland

Local Organizing Committee

Dimitrios Papanikas	University of Piraeus, Greece
Kleanthis Dellios	University of Piraeus, Greece
Theodoros Ntouskas	University of Piraeus, Greece
Evelina Peristeri	SigularLogic S.A., Greece

Program Committee Chair

Christos Douligeris	University of Piraeus, Greece

Program Committee

Esma Aimeur	University of Montreal, Canada
Joao Paulo Almeida	Federal University of Espírito Santo, Brazil
Hernán Astudillo	Universidad Técnica Federico Santa María, Chile

Evandro Baccarin	University of Londrina, Brazil
Khalid Benali	Loria - Nancy Université, France
Salima Benbernou	Université Paris Descartes, France
Djamal Benslimane	Lyon 1 University, France
Sami Bhiri	National University of Ireland, Galway, Ireland
Markus Bick	ESCP-EAP Europaische Wirtschaftshochschule Berlin, Germany
Melanie Bicking	University of Koblenz-Landau, Germany
Gérôme Canals	Loria - Nancy University, France
François Charou	Henri Poincaré University of Nancy 1 - LORIA, France
Vassilis Chrissikopoulos	Ionian University, Greece
Ioanna Constantiou	Copenhagen Business School, Denmark
Bruno Defude	Telecom SudParis, France
Alicia Diaz	Universidad Nacional de La Plata, Argentina
Toumani Farouk	Blaise Pascal University, Clermont-Ferrand, France
Fernand Feltz	Centre de Recherche Public - Gabriel Lippmann, Luxembourg
Simone Fischer-Huebner	Karlstad University, Sweden
Emmanouil Georgakakis	University of Piraeus, Greece
George Giaglis	Athens University of Economics and Business, Greece
Claude Godart	LORIA, France
Americo Nobre Goncalves Ferreira Amorim	Faculdade Integrada do Recife, Brazil
Adam Grzech	Wroclaw University of Technology, Poland
Mohand-Said Hacid	Université Claude Bernard Lyon 1 - UCBL, France
Tomasz Janowski	UNU-IIST, Macau
Marijn Janssen	Delft University of Technology, The Netherlands
Atsushi Kanai	HOSEI University, Japan
Sokratis Katsikas	University of Piraeus, Greece
Ranjan Kini	Indiana University Northwest, USA
Angeliki Kokkinaki	University of Nicosia, Cyprus
Panayiotis Kotzanikolaou	University of Piraeus, Greece
Dennis Kundisch	University of Paderborn, Germany
George Lekakos	Athens University of Economic and Business, Greece
Hongxiu Li	Turku School of Economics, Finland
Jose Machado	Universidade do Minho, Portugal
Tiziana Margaria	University of Potsdam, Germany
Haris Mouratidis	University of East London, UK
Jose Neves	Universidade do Minho, Portugal

Andreas Oberweis	Universität Karlsruhe, Germany
Harri Oinas-Kukkonen	University of Oulu, Finland
Boissier Olivier	ENS Mines Saint-Etienne, France
Hervé Panetto	Nancy University, CNRS, France
Spyridon Papastergiou	University of Piraeus, Greece
Olivier Perrin	LORIA, France
Riedl Reinhard	Bern University of Applied Sciences, Germany
Florence Sedes	I.R.I.T. Paul Sabatier University, France
Santosh Shrivastava	Newcastle University, UK
Paul Spirakis	University of Patras, Greece
Zhaohao Sun	University of Ballarat, Australia
Reima Vesa Suomi	Turku School of Economics, Finland
Dimitrios Vergados	University of Piraeus, Greece
Herve Verjus	Université de Savoie - LISTIC - Polytech'Savoie, France
Maria Virvou	University of Piraeus, Greece
Hans Weigand	Tilburg University, The Netherlands
Robert Woitsch	BOC, Austria
Hiroshi Yoshiura	The University of Electro-Communications, Japan
Steffen Zimmermann	University of Innsbruck, Austria

Co-editors

Christos Douligeris	University of Piraeus, Greece
Nineta Polemi	University of Piraeus, Greece
Athanasios Karantjias	SingularLogic, S.A., Greece
Winfried Lamersdorf	University of Hamburg, Germany

Table of Contents

Protocols, Regulation and Social Networking

Adoption Issues in e/m-Services

New Services Adoption and Ecological Behaviour

Knowledge Management and Business Processes

Management, Policies and Technologies in e/m-Services

Technology Regulation 2.0?

Andreas Mitrakas

Abstract. This paper discusses future ways to regulate technology effectively in an ever accelerating speed of technological development. The current rate of response of legislation to technological evolution is unlikely to allow law keeping up pace with and responding to the needs of society for technology regulation. This is likely to create a rift between on one hand the perceived need of society for technology regulation and on the other, the ability of the legislative to deliver in line with expectations. While the demand for more technology regulation in unlikely to subside, self -regulatory means and tools are likely to increase in prominence, especially if combined with technology in terms of automation. Societal value could be developed in a way that allows legislation to retain its important role in meeting the need for greater control of the evolution of technology. The approach proposed in this paper could be of use to those involved in the governance, service provision or use of technology regulation.

Keywords: Information technology, strategy, regulation, cost benefit analysis, standards, policy, soft law, contracts.

1 Introduction

Technology, which is the outcome of creative and innovative mental processes, requires regulation to become usable in society and the economy. Regulation in many ways has provided the safe haven for human ingenuity to thrive, protected, and expand further. In spite of its shortcomings, intellectual property regulation, has provided the ground upon which technology advancements of the industrial and post-industrial ages have come about. In the information age, technology has been driven by private sector initiatives, that seek to penetrate global markets or markets that present comparable features across multiple jurisdictions. Regulating for technology has challenged and often deluded operators, users and legislators alike due to the complexities involved in grasping the often far-reaching consequences of technological innovation. There is an obvious lag between what the legislator sees and believes that technology can do and what technology actually does as it is put to work in a performance driven economic environment. This is no easy rift to bridge and in response Law has sought for decades to restore balance by having Courts interpreting apparently obsolete legislation to render it fit for new purposes. There is no shortage in examples of technology regulation that become obsolete as a result of innovation. This paper reviews some examples of innovations that are credited for the state of art of technology services it then seeks to present an approach on how regulation could evolve to accommodate the emerging needs of technology regulation.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 1–12, 2013.

2 Background

Future is determined by such elements as, nascent and emerging technologies the impact of which on society and the economy is not quite appreciated at the time of their emergence. The impact of technology is related to the interaction of emerging technologies with existing ones, the standing need to regulate and the need to organise regulation in a meaningful way that maximises societal benefits from the introduction of technology. However as the pace of technology picks up, the regulation of technology has to follow suit; and this remains a challenge that has yet to be tackled. Current legislative review cycles last long and it is not uncommon that from one "generation" of regulation to the next 20 years lapse; this lag has consequences, as the use of technology under innovative business models eludes the judicial that has to resort on interpretations to accommodate change. The far reaching consequences of technology are often not understood well enough when a technology is launched or when a technology is put under legislative scrutiny. When applying "dated" laws on new use cases sometimes more space for interpretation is permitted than what is actually needed.[1]

The past two decades have allowed for the fusion of information technology with an array of other technologies accelerating the pace of technology expansion to unprecedented levels. In 1965, Moore's law described a pattern for the exponential growth in complexity and in capacity of integrated semiconductor circuits. Moravec extrapolated Moore's Law to future forms of technology, such as robotic agents.[2,3] Kurzweil suggests that the pace of technological expansion might accelerate further and become the key feature of technological development as early as in 10-15 years.[4]

In the 1960s and through to the 1980s the promise of expert systems gave new impetus to what machines could accomplish to replace humans in carrying out intelligent tasks. Realising that law had a formal as well as an empirical set of attributes that cannot be represented in the same way brought about a break through. Empirical knowledge was brought under the spotlight and it was added as a feature to

[1] See, Vernor v. Autodesk, Inc. (555 F. Supp. 2d 1164), that addressed the applicability of the first-sale doctrine to software sold under the terms of "shrinkwrap licensing." The First Instance Court ruled that the transfer of software to the first acquirer that bore the characteristics of a sale gave right to reselling the software under the first-sale doctrine. In appeal, the US Court of Appeals for the Ninth Circuit, (decision of 10/09/2010), reversed the first sale doctrine ruling remanding for further proceedings on the misuse of copyright claim. In a more recent development the Librarian of the Library of the Congress applied this principle on the jail breaking practice concerning smart phones, rendering it illegal under the Digital Millenium Copyright Act (DMCA - Pub. L. 105-304, Stat. 112 Stat. 2860 (1998)), unless the operator agrees to it. This is a development that the legislator of the DMCA, that aimned at protecting the intellectual property rights of content providers, could hardly foresee in 1998.

[2] See, Moore, G.E. (1965). Cramming more components onto integrated circuits. Electronics, 38:8.

[3] Moravec, H. (1998). Robot: Mere Machine to Transcendent Mind, OUP.

[4] Kurzweil, R., (2006). The Singularity is Near. Gerald Duckworth & Co Ltd.

the requirement of formalised knowledge when seeking to automate legal procedures as it was the case of legal expert systems for example.[5]

Automation in business transactions became yet another frontier to explore. The evolution and widespread use of formalised data structures, i.e. electronic data interchange in select business-to-business processes made automation and exchange of structured business data widespread leading to significant efficiencies marked in terms of costs and working methods. The opening up of telecommunications networks led to the gradual but increasingly more accelerated later pace to develop new products and services. The emergence of the commercialised World Wide Web allowed for further inroads in terms of automation and decision support and commoditisation of information. Intelligent agents have been able to collect preferences, make suggestions and support decisions made by consumers in large applications environments for electronic commerce. Large scale reputation systems profiled users enhancing the sense of trust that consumers and business users alike can show in a transaction. In a case regarding transactions with private and public bureaucracies, an intelligent agent would fetch information from various sources, i.e. libraries to present it to the requestor according to a predefined set of presentation rules and help the requestor accomplish an organisational goal.

On the regulatory side the need to protect personal data from undue breaches led to the adoption of data protection legislation. While the first concrete piece of European legislation on data protection was voted in 1995, in January 2012, the European Commission proposed a sweeping reform of the EU data protection rules of 1995. The 1995 rules had been interpreted differently by the 27 member countries, which led to inconsistent enforcement. The new proposal suggests a single law that will apply to all members of the EU. This proposal is now under legislative scrutiny by the European Council and the European Parliament before becoming a final law in 2015.

3 Technologies of the Day after Tomorrow

Various technologies are likely to play a key role in the years to come, and shake the current technology pattern by providing a breakthrough. Obviously the technology landscape will not only be influenced by technology. It is likely, however that a small number of them might change pre-determined patterns, shifting social and regulatory interactions related to these technologies. The impact of these technologies is such that unless grasped early enough by the legislator with a view to re-evaluate its regulatory strategy, they are going to be disruptive in a way that regulating them becomes dysfunctional.

Big Data

Big data is extremely large sets of data related to consumer behaviour, social network posts, geo-tagging, sensor outputs, and other types of tags and meta-tags that

[5] Leith, P. (2010). The Rise and fall of the legal expert system. European Journal of Law and Technology, 1:1

associate human behaviour with data.[6] On the outset big data is the output of devices and sensors that consumers, organisations and devices use across the globe to go about their lives and business. Data is generated due to human or machine activity; it is recorded by means of information systems that are hosted typically in large data centres on the cloud. Big data is an extended representation of human activity and behaviour in all walks of digital life and beyond. Leveraging on big data broadens the ability of individuals and organisations to interact, as well as monitor, control and evaluate such interactions. It is likely that big data will pose such challenges as access to data, privacy, data formatting and challenges for security.

Agent Driven Identification and Authentication Technologies

In the draft European Regulation on electronic identification and authentication services legal persons are afforded the ability to sign independently of the natural persons involved in their management, like a general manager representing the organisation for example.[7] In view of the explicit permission for legal persons to sign an additional extension could be considered to legally permit intelligent agents to sign for the purpose of non-repudiation. If intelligent agents are not explicitly permitted to sign with an electronic signature and bind for example a natural person, would it be within reach to incorporate an intelligent agent and have it sign on behalf of the incorporating company, which as legal person is allowed to sign on its own.

This is a theoretical argument for the time being of course, but in the course of the next few years leading to implementation of this Proposal, signatures for intelligent agents are likely to become a necessity.

As technologies become ubiquitous in the next 2-5 years that is the period in which this Proposal will be voted and its implementation will be launched it is likely that the next legislative action in the area of electronic identification and authentication will be in a period shorter that the 15 year that that Directive 1999/93/EC claimed from launch as a proposal till sunset. It is likely that legal capacity of agents to conclude legally binding acts might be challenged, especially in view of possible combinations with other provisions such as incorporating an intelligent agent. Further to the capacity to act, other areas might include the validity of a legal act signed by an agent, the evidence required to prove an act and the value of such evidence. Information security and privacy aspects are likely to become of importance too as accessing the data storage of the intelligence agent might reveal patterns related to natural persons.

Coming to Terms with the Law of Accelerating Returns

According to Kurzweil "an analysis of the history of technology shows that technological change is exponential, contrary to the common-sense "intuitive linear" view. So we won't experience 100 years of progress in the 21st century — it will be more like 20,000 years of progress (at today's rate). The "returns," such as chip speed and cost-effectiveness, also increase exponentially. There's even exponential growth

[6] Johnson B.D. (2012), Big data; Still to come? The Futurist. 46:4.

[7] Proposal for a Regulation of the European Parliament and of the Council on Electronic identification and trust services for electronic transactions in the internal market (COM(2012) 238).

in the rate of exponential growth.'' Within a few decades, Kurzweil claims, machine intelligence will surpass human intelligence allowing a merger of biological and non-biological intelligence."[8] In an era of unanticipated technological developments, assessing the capacity to transact of new types of human intelligence might be required.

4 Towards Regulating Potentially Disruptive Technologies

Even before the notion of exponential technological growth was introduced, the impact of technology had been typically disruptive due to the new and sometimes uncalculated elements that it introduced in society, as it was applied. Disruptions might not be visible to the unaware observer from one stage to the next however the impact technology has other societal expressions has a dramatic effect. New business ideas that rely on technology appear and old ones that are based on obsolete or concurrent technologies become extinct. It is quite challenging to have new technologies rooting as viable business ideas especially as they seek to compete with other established ones that in terms of business cycle they resonate well with the users. Some aspects of the economic perspective of regulation, that seeks to address the governance, social and economic order beyond regulating rights and obligations of parties involved is further presented.

The efficiency of a proposed regulatory solution can be analysed on the basis of the Coase theorem that provides us with the notion of what efficiency means when making an agreement.[9] Coase suggests that a contractual solution must take into account three factors: transaction costs, the efficiency of the outcome and the legal framework.[10] By considering these variables a solution can be reached. The efficiency of a contract also depends upon factors such as the social context, the transaction costs and the legal reality of the environment in which the legal solution is applied. If the cost of obtaining information in order to draw up a service agreement concerning information security services is too high, other solutions should be sought instead.[11] If we apply this theorem in the regulatory area for technology we can consider that regulation is preferred when unfettered access to information is available; i.e. access to access to all possible future uses of technology is clearly described, sufficiently understood and anticipated before the regulation is proposed.

In a mature technology environment, where a broad selection of technologies and regulatory instruments are available, several factors influence the final regulatory outcome. These factors include negotiation power of the key actors, the expertise they have, access to resources, the timetable to fulfil the desired level of implementation, etc.

---

[8] See, `http://www.kurzweilai.net/the-law-of-accelerating-returns` (visited on 30/01/2013).

[9] Posner, R., (2003) *Economic analysis of Law*, New York: Aspen Publishers.

[10] Mitchell Polinsky, A., Shavell, S. (2010). *The Uneasy Case for Product Liability*, Harvard Law Review.

[11] A. Mitchell Polinsky, S. Shavell, *The Uneasy Case for Product Liability*, forthcoming, Harvard Law Review (forthcoming), 2010, pp. 31-34.

In general to decide on whether regulation is required the regulatory burden (RB) of legislating, monitoring compliance and enforcing should be calculated and pondered against the perceived market distortion costs (MDC) due to the associated with the lack of regulation. The relationship could be presented as follows:

$$RB<MDC$$

In regulating technology, law has experienced long return cycles that allowed for consolidated views to emerge, consensus with regard the proposed legislative arrangement to emerge, those who could potentially disagree to drop out of grace or be taken over or having their views represented by others who sought the proposed legislative solution. There are numerous examples of legislation that even in the more recent past of rapid technological growth required return cycles of fifteen years to pass from one generation of legislation to the next.

There is little doubt that the role of technology is likely to expand further and penetrate yet new levels of human activity in the society. Further regulation will remain a requirement. However in view of the accelerating pace of development it is doubtful that in a few years from now there will be sufficient time to effect regulatory evolution the way we are used to do. Those long review cycles are likely to be reduced to a few months or weeks, reducing the capability of the legislator to respond, or deeply understand the implications of both technology and regulation. This alone would not be the problem of course; the issue is related with the societal impact that the inability of the legislative process to bring about regulatory impact will have. While large swaths of technological activity would remain under regulated or outside regulatory scope the impact of the lack of regulation would be stark denying business and citizens protection from abuse and the ability to enforce rights.

5 Seeking to Regulate in the Age of Exponential Growth

The pace of regulation is not likely to take a recess; however the way regulation is promulgated is likely to face new challenges with the advent of ever shorter technology cycles that bring disruptions along with them.

The Accelerating Disruption Cycles Paradox

It is also likely that the need to govern the cross-border impact of technology as well as the countering of threats to society that are brought about by technology will also pick up pace. Therefore a paradox is likely to emerge being that from an ever narrower window of opportunity due to the accelerating disruption cycles more regulation will be demanded; this is likely to impact future regulatory strategies. This situation is unlikely to be sustained, should the legislator resort on currently available methods and means. A combination of efficient means is likely to emerge to facilitate the legislator in bringing about regulatory change and following the pace of technological development.

The remainder of this paper reviews some models that could bear fruit when seeking to invoke self-regulatory elements in information technology without necessarily resorting to new legislative instruments. While the legislative impetus on information technology is likely to pick up further in the future, concurrent approaches could help reducing the effect of the long legislative process on trade partners.

Commoditising Valuation Models

The valuation of goods and services is likely to be more in demand. In case of civil wrongdoing, like when a contract is breached it is essential to reckon what has been the value of the wrongful act done. The answer very much depends on the value that the goods or services had at the time the wrongful act occurred. There is a direct relationship between the goods or services in question and the markets in which these values are measured. If A is involved in a car accident with the 10-year old car that belongs to B, clearly A owes damages to B, however those damages will be discounted in a way that reflects the net present value of the car at the time of the accident, rather than the value of the car when new. Is B is unhappy with its valuation because of the emotional value that this good carries for him, additional elements might have to be inserted in the calculation to satisfy B (the answer cannot be to damage something of emotional value to A); this of course might be a difficult calculation, as the only objective way to value that car is to base it on a market calculation.[12]

If the goods at risk are intangible, the example might have a different outcome. Home videos on a memory stick have to be transferred to different media in a shop; the clerk pays attention but somehow the home videos are lost. If the possibility to lose a video is 1/1000 then how would these videos be valuated? The photo shop clerk offered to deliver the videos in a steel box and never lose them against an extra fee of 100 Euros. The client is faced with the dilemma of insuring against a non-valuated item. Suppose the client decides to pay the premium then the valuation he attaches to it is 100 * 1000 = 100 000 Euros.

But how can data owners assess their potential damages by themselves, in information security? Although this might sound unusual, it is not; insured parties typically carry out an implicit risk assessment as well as a self-assessment of the value of the insured goods or services before they turn to an insurance broker to get a quotation. A self-assessment of the possible value of the information that is under threat is a first but critical step down the road of securing information. And this is a step that the information owner who is the user of the information security services has to take. There is a discreet benefit in following this approach, as the information owner is likely to have to take two dimensions of the problem into account: the information owner will want to state a low value to keep his insurance premiums down but if he states too low a value, then the value of his intangibles might be deemed to be too low for his overall business valuation.

[12] Ward Farnsworth, *The Legal Analyst: A Toolkit for Thinking about the Law*, University Of Chicago Press, 2007, location 4364.

A different situation arises from a combination of data that serves invariably personal and professional proposes; such is the case of a home office where a professional stores home videos and company data alike. What would the value of the loss of data be in this case? While information security measures can indeed by employed to protect private and professional data alike, it is highly unlikely that the valuation of the loss of data will be identical for both categories of data. When seeking benchmarks, the "willingness to pay" comes forward as a proper measurement to benchmark the situation. Suppose that an airbag costs 100 Euro and that every 10 000 purchases an airbag gets to save one life; in this case 1 million is spent to save one life, which sets the benchmark for a price tag on a life to 1 million. The purchaser is protected from this 1/10 000 chance to lose his life in an accident involving the use of an airbag. If airbags were priced at 1000 Euro and people stopped buying them this would mean that the price on life is below the 10 million mark.

In technology the impact of the loss of information has to be valuated by the client as we are talking about intangibles that typically carry subjective value that might only be of importance to the owner of that information. There is information of course that serves purposes that can be determined in an objective manner, like it is the case of tax related data held by the tax payer. As cloud service providers proliferate their services it is likely that increasingly more data will flow from proprietary data centres that operate at arm's length to cloud based services. If the jurisdiction of protecting data in transit or at rest has not been determined it is likely that the client might find herself in a challenging situation if data is wrongfully handled, tampered with or outrightly lost.

Valuation models for data are likely to be further developed for the purpose of facilitating trade partners and consumers alike in a way that allows them to better leverage on new business models.

Cost Benefit Analysis

What are the appropriate organisational arrangements concerning risk quantification that allow the data owner to properly value and assess the value of the data under her control? In the air safety or maritime safety areas, government agencies set the stage of standards and metrics required to meet the safety expectations of the industry. This is a top down approach that has evolved over time and it ensures that benchmark requirements are in place and they are met. An alternative approach is to take a contingent valuation whereby the target group is surveyed as to how much they are prepared to pay in order to avoid a certain result from happening.

Cost benefit assesses the total expected costs against the total expected benefits to obtain the most profitable option. Quantified benefits and costs are adjusted in order to calculate their present value. Economic efficiency is taken as a criterion to assess the relative merits of two situations and then decide which one is the best able to maximize social welfare, after taking into account social costs and benefits. Cost-benefit analysis is developed as a subject in order to be a practical guide to social decision-making by evaluating social costs and benefits with private costs and

benefits. It is often the case that the project should go ahead while the benefits to society as a whole from the project are greater than the benefits of any individual pertaining to this project. The efficiency effect of a project is the difference between costs and benefits. In order to calculate the social costs and benefits it is necessary to consider also private costs and benefits as well as externalities.[13] A cost-benefit analysis allows to objectively assessing the benefits deriving from a certain action. However the criteria used in a cost-benefit analysis are typically rigged by the subjective views of the party that carries out such analysis. Cost-benefit analysis refers to assessing, the case for a project, and program or policy proposal for the purpose of supporting decision making.

The highlights on this area include certain aspects that provide assurance to the parties involved. Self-regulatory efforts that can take the form of invariably seeking compliance with the policies and technical standards in the application area; setting out a contractual framework that involves all parties in an information security framework regardless of them being service providers beneficiaries or relying parties etc.

In terms of quantifying liability as such the aspects that can be considered include the calculation of a cost benefit co-efficient that can be used to assess the value extracted from the information security measures employed. Additionally a market valuation should be made available that sets out the value that data represents for the data owner vis-a-vis its market usability; low value personal data can be set with the lowest value, while high value professional data can be assigned with a high value.

Most importantly an assessment of what data represents for the information owner stands at the centre of this approach that aims at determining the liability emanating from loss of data in case of an information security breach.

Standards for Self-regulation

Information technology woes are often addressed by means of voluntary frameworks self-imposed by the trade partners themselves. These frameworks include policies and agreements that aim at setting up the conditions for information security safeguards within an organisation, or in transaction frameworks. At a bilateral level, the parties use service level agreements to specify the quality service they seek from their provider and ensure availability rates for their applications. Parties might set up security frameworks, which are activated by means of subscriber agreements executed individually. In this later example the service can be a generic one that does not necessarily allow for a high degree of customisation. Information security assurance can be provided for by adhering to international standards. Regardless of the form information takes, or means by which it is shared or stored, it should always be appropriately protected.

Standardisation in the EU has gradually assumed an approach that contains covers up for limitations emanating from a strict legislative process that can be seen as too

[13] Weixiao Wei, (2008) *ISPs' Indirect Copyright Liability Regime: An Economic Efficient Liability Regime for Online Copyright Protection Shaped by Internet Technology*, Bileta Conference Proceedings.

narrow to address business and society needs.[14] Technical standards, thus, provide a layer of "soft law" that deviates from a strict legislative approach; in promulgating "soft law" standards, the industry often has a leading role.[15]

Liability Assessment Tools

A model is further presented hereunder to support the process of assessing liability. The application area of this very model is limited and it presents a calculation model that aims at quantifying liability. This conceptual Information Security Liability Assessment (ISLA) tool aims at easing typical shortcomings associated with the assessment of the value of information in a given operational environment that pose discreet shortcomings to the ability of risk takers to determine and apportion their liability exposure.[16]

ISLA: Information Security Liability Assessor	
Assessment criteria	**Representation value**
Own valuation of data	10
Market valuation (market risk co-efficient)	0…+3
Cost benefit	0...+2
Policy (audited and compliant)	+1
Contract	+2
Standards	+1
Worst case	-3 (variable)

The model works as follows: the information owner has to determine the value of the data (DV) that he seeks to protect by means of information security measures. A market value (MV) factor represents the social impact of the data loss. In this case a scale form 0-3 is permitted where 0 denotes strictly personal, no value data and 3 represents the high grade professional data to be protected.

A cost benefit (CB) analysis indicates whether there is need to outsource the task of information security or it makes better sense to carry it out as an in sourced one.

[14] Standardisation policy in the EU dates to the *Cassis de Dijon* case in which the European Court of Justice ruled that a product meeting the requirements of one member state should be legally made available in another; allowing the emergence of mutual recognition of technical standards as a matter of significant interest in the EU internal market (ECJ 120/78 of 20/02/79).

[15] Senden, L., (2002). Soft law in the European Community Law. Hart Publishing.

[16] Mitrakas A., (2011) Information security liability: an assessment model, in P. Kleve, P., Noortwijk van, K., Something bigger than yourself: Essays in honour of Prof. R. V. De Mulder, Rotterdam: Erasmus University Rotterdam; See also, Mitrakas, A., (2011) Assessing liability arising from information security breaches in data privacy, International Data Protection Law Journal, Vol. 1, Issue 2, Oxford: Oxford University Press.

Acceptable values range from 0 to +2 to denote the efficiency rate of a solution, with low figures referring to own solutions and a higher figure leading to high risk that are mitigated with high protection alternatives.

There are 3 conditions that could ease things up if properly implemented, and these include, a strict and adhered to policy framework (P), a contractual framework (C) and the implementation of technical standards (S).

It is advisable to remove points if a "worst-case scenario" (WS) has to be provided for, like in case of high risk operations. We then derive the following formula that helps assessing the estimated liability (L) exposure.

$$L=((DV*MV)/CB)*((P+C+S)-WS)$$

The higher the result the more valuable the data asset is assessed, which boosts in return the liability in case of an information security breach. The overall results can be utilised in an operational context to determine the liability of an information security breach.

Clearly this model does not provide an absolute instrument to determine liability and variations or interpretations might be necessary in some cases. Isla can, however be used in a coherent manner when seeking to determine the liability of service providers or users of information security services.

6 Back to the (legislative) Future?

When seeking to lay out legal strategy for future legislative action that meets new technology challenges, it is extremely hard to make predictions of course. The legislator has the ability to determine the way it passes legislation anyway, hence this article rather than making predictions seeks to present a snapshot of what is possible. The challenge posed by technology ahead is associated with rapid change. Long legislative cycles are unlikely to survive as a way of addressing the regulatory requirements of technology. Nevertheless, the demand for legislation to regulate technology is likely to increase. Legislation is thus likely to have to be more geared towards a two-tier model, where at the top tier, a broad brush approach based on principles and governance will be adopted. At this level the high level issues such as principles, a governance model and methods are addressed. At a lower level the implementation will be organised around the needs for benefiting from new technologies, quick liability simulations or rapid extra-judicial settlements, by means of automated agents. To support regulatory needs methods that draw their origin from other disciplines are also likely to emerge. At the European level standardisation has been the typical has example that has been leveraged upon to deliver regulation according to predefined models supported by the parties upon which regulation is directed [17]. Other models might include cost benefit analysis, risk assessments and more technology oriented ones such as liability valuation models.

[17] Mitrakas A., and Eecke van P. (2005). Commentary on Directive 1999/93 on a Community framework for electronic signatures, in Buellesbach, A., Poullet, Y., Prins J.E.J. (eds.). Alphen aan de Rijn: Concise European IT Law, Kluwer Law International.

7 Conclusions

The impact of technology on society has led to numerous efforts to address technology governance and transaction risks by means of legislation. The pace of legislation with long review cycles is likely to be affected in the light of technology convergence and the emergence of new technologies that further accelerate the technology life cycle. In the near future it is likely that the legislative process will have to pick up pace to match that of technology. This is a challenging matter. Alternative strategies can be developed to allow for the regulation of technology by reducing that the legislative process has. In this paper we presented an array of instruments that can be seen as a way to regulate technology and user behaviour in a world of accelerating technology cycles.

Note

This paper expresses the author's personal views only and not those of any other party including his employer in any way whatsoever. This paper cannot be associated with any deliverable, report or opinion of the author's employer.

References

Farnsworth, W.: The Legal Analyst: A Toolkit for Thinking about the Law. University of Chicago Press, Chicago (2007)

Johnson, B.D.: Big data; Still to come? The Futurist 46(4) (2012)

Kurzweil, R.: The Singularity is Near. Gerald Duckworth & Co. Ltd. (2006)

Leith, P.: The Rise and fall of the legal expert system. European Journal of Law and Technology 1(1) (2010)

Mitchell Polinsky, A., Shavell, S.: The Uneasy Case for Product Liability. Harvard Law Review (2010)

Moore, G.E.: Cramming more components onto integrated circuits. Electronics 38(8) (1965)

Mitrakas, A.: Information security liability: an assessment model. In: Kleve, P., Noortwijk van, P. (eds.) Something Bigger than Yourself: Essays in Honour of Prof. R. V. De Mulder. Erasmus University Rotterdam, Rotterdam (2011)

Mitrakas, A.: Assessing liability arising from information security breaches in data privacy. International Data Protection Law Journal 1(2) (2011)

Mitrakas, A., Eecke van, P.: Commentary on Directive 1999/93 on a Community framework for electronic signatures. In: Buellesbach, A., Poullet, Y., Prins, J.E.J. (eds.) Alphen aan de Rijn: Concise European IT Law. Kluwer Law International (2005)

Moravec, H.: Robot: Mere Machine to Transcendent Mind. OUP (1998)

Posner, R.: Economic analysis of Law. Aspen Publishers, New York (2003)

Senden, L.: Soft law in the European Community Law. Hart Publishing, Oxford (2001)

Wei, W.: ISPs' Indirect Copyright Liability Regime: An Economic Efficient Liability Regime for Online Copyright Protection Shaped by Internet Technology. In: Bileta Conference Proceedings (2008)

Order Effects in Observations of Stated and Revealed Privacy Preferences

Christian Flender

Institute of Computer Science and Social Studies,
University of Freiburg Freiburg, Germany
christian.flender@iig.uni-freiburg.de

Abstract. Many Internet services rely on consumers disclosing their personal data. Despite heavy usage and wide acceptance of services like Online Social Networks, doubts about sustainability of trusted relationships remain. Surveying consumers about their preferences reveals severe concerns about the fate of their personal data. In stark contrast to privacy concerns (stated preferences), however, consumers generously disclose personal data in exchange for free Internet services (revealed preferences). It has been argued that individuals experience dissonant states in privacy decision making. The tension between stated and revealed preferences is eliminated with the decision made in order to reduce discomfort. This paper proposes a survey design to determine 1) order effects as indicators for dissonant states in privacy decision making, and 2) the degree of experienced tension between stated and revealed preferences. Observations of data valuation and disclosure behavior are dissonant if they do not commute, i.e. disclosing data prior to valuating privacy does not equal privacy valuation before data disclosure. Determining the degree of dissonance in privacy decision making is expected to inform the design of transparency mechanisms to influence experienced dissonance between stated and revealed privacy preferences.

Keywords: Privacy Decision Making, Cognitive Dissonance, TET 2.0.

1 Introduction

With the global acceptance of Internet services dealing with personal data, generous disclosure behavior of individuals has posed pressing research problems. One of the biggest issues at present is the challenge to establish trust among all the Web participants involved in order to sustain the benefits of digital life for individuals and society at large. Measuring trust or the lack of it, however, is a difficult exercise, last but not least due to paradoxical preferences stated and revealed by Web users. What came to be known as the privacy paradox measures the contradiction between high valuation of privacy and generous disclosure of personal data [1]. Many attempts to explain this discrepancy have been put forth from psychological, sociological, and economic perspectives. Some authors argue that individuals attribute a higher value to the benefits of Web services arriving sooner in time compared to the associated costs and

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 13–24, 2013.

long-term risks in the far future [2]. Others share the view that peer group pressure imprisons Web users in Online Social Networks effectively rendering opting out behavior and avoidance of data disclosure undesirable [3]. Economically, it has been argued that Web users have information deficits about the way their data is used and valued on data markets which makes them careless about privacy matters [4].

More recently, privacy decision making has been put under the light of cognitive dissonance theory [5]. The theory of cognitive dissonance attempts to explain the behavioral phenomenon of a person experiencing two conflicting cognitions (behaviors, ideas, beliefs, or attitudes) [6]. People reduce or eliminate the tension arising from the dissonance by either changing cognitions or adding new cognitive elements in order to reduce discomfort. Subjects in dissonant situations experience an inconsistent state of mind and with the decision made they rationalize their choice in order to re-establish a consistent state. For instance, a smoker who knows or learns about negative health effects, either stops smoking, changes his attitudes towards smoking, or adds explanations about smoking to reduce dissonance, e.g. car driving is more dangerous than smoking. Likewise, Web users are supposed to experience two conflicting situations. On the one hand, they appreciate the right to be left alone in private. But when confronted with attractive Web services they change their attitude and generously disclose sensitive information about their personality and private life. Cognitive dissonance seems to have an impact upon privacy decision making and raises the following research questions:

1. Do Web users experience cognitive dissonance in privacy decision making?
2. What is the degree of cognitive dissonance in privacy decision making?
3. Can Transparency Enhancing Technology (TET) influence cognitive dissonance and trigger change?

The objective of this paper is to propose a design to measure cognitive dissonance in privacy decision making and pave the way for further studies on the effectiveness of Transparency Enhancing Technology (TET). TETs like privacy statements or dashboards for personal data management are meant to enable Web users to exercise their right to informational self-determination, i.e. their right to control, edit, manage, and delete personal information and decide when, how, and to what extent it is communicated to others [7]. At present, TETs are hardly used. Offering to Web users incentives thereby influencing dissonant behavior may inform TET 2.0, a new generation of transparency mechanisms changing user behavior towards trusted relationships based upon their free self-determination.

The method of research presented in this paper follows the design science paradigm [8]. A survey design is developed and implemented on the basis of methodological considerations on how to operationalize the measurement of cognitive dissonance in privacy decision making. Drawing from Flender and Müller (2012), privacy decision making is inconsistent or dissonant if observations of stated and revealed privacy preferences are not interchangeable, i.e. they do not commute [5]. The design developed in the following accounts for order effects in measurements of privacy preferences.

2 Order Effects Indicate Dissonance in Privacy Decisions

The baseline of many privacy research problems is the paradoxical observation of generous data disclosure and high appreciation for privacy [9,10,11,12]. When people reflect upon the value and sensitivity of their personal information most of them are deeply concerned about the fate of the traces they leave on the Internet. Contrary to stated preferences, however, most people reveal a low valuation of their personal data when it comes to eased communication and cooperation on the Internet, comfortable online shopping, or reciprocal attention spending in Online Social Networks. A recent explanation of the privacy paradox was developed in [5] and draws from cognitive dissonance theory. According to preference (type) indeterminacy in privacy decisions, the paradoxical nature of privacy decisions stems from the fact that states of decision outcomes are constructed with the decision made and thus do not exist prior to their construction. Moreover, the state of a decision outcome depends on the sequential order of successive measurements of privacy preferences. If sequential observations of stated and revealed privacy preferences interfere, order effects apply and indicate dissonance.

Mathematically, two conflicting privacy decision situations are modeled as two observables, or properties. The outcome state of a sequence of successive privacy decisions is determined with the decision made but not prior to it. For instance, measuring "high valuation of privacy" and observing "generous data disclosure" are two states of two privacy decision situations A (stated preference) and B (revealed preference). By means of a numerical example representative of the privacy paradox, it was shown that the sequences AB and BA do not commute, i.e. $AB_p \neq BA_p$ where p is a state after two sequential decision situations A and B with the potential states *a1*) high valuation and *a2*) low valuation for A, and *b1*) data disclosure and *b2*) data concealment for B. Prior to an actual decision, individuals reside in a dissonant state of potentiality, i.e. a linear superposition a1 + a2 and b1 + b2 respectively. With the decision made the tension between A and B reduces the dissonant state to p which marks an opinion change, a behavioral change, or rationalization. It was shown that A and B cannot be measured simultaneously, i.e. they are incompatible or do not commute, by proving the probability of data disclosure after data valuation to be lower than the probability of data disclosure before data valuation. In other words, order effects in sequences of successive measurements of stated and revealed preferences model the dissonant situation of individuals experiencing a tension between two conflicting though not yet determined cognitions.

Theoretically, cognitive dissonance offers a new explanation of why people disclose data they way they do by arguing that privacy preferences are indeterminate and dissonant up to the point in time a decision is made. Stating preferences and revealing preferences about personal data are not interchangeable. Instead observations of stated and revealed preferences interfere, i.e. order effects apply when stating preferences before revealing preferences and vice versa.

The next section discusses studies for determining privacy preferences empirically. There is a multitude of methods for measuring Willingness-to-Pay (WTP) for public and private goods. Each method only represents the attempt to come as close as possible to the truth [13]. There is not one best method. Rather appropriateness of a method depends upon validity and feasibility with regard to a given research problem. Validity of a method for measuring cognitive dissonance in privacy decision making has to account for the context-dependent nature of privacy preferences. Preferences are not out there readily determined. In fact, preferences are never revealed but constructed with the decision made. Two basic methodological assumptions derive from the discussion so far. First, the design for measuring preferences abstains from naive realism, or the claim that true preferences are out there readily determined. Second, the design assumes that observing order effects is a necessary condition for dissonant privacy decisions.

3 Observations of Privacy Preferences

To determine preferences of individuals about products and services Willingness-to-Pay (WTP) measures the value someone attributes to a material or ideal good. Generally, direct and indirect WTP methods are distinguished [14]. The former ask for the price someone would be willing to pay for a good. This could be the maximum price someone is willing to pay or the minimum price someone would charge for a good. Indirect methods survey participants about the legitimacy of a given price for the product in question. The problem with direct methods is the variability of preferred prices as individuals may differ substantially in their valuation of goods. Asking for the legitimacy of a given price bears the problem that answers provide less information about the true value individuals attribute to the product. Generally, WTP methods strive for incentive compatibility, i.e. subjects ought to reveal truthfully any private information about their willingness to pay for a good without taking strategic considerations into account, e.g. concealing private information now may bring bargaining advantages in the future. Thus incentive compatibility presupposes preferences to be out there readily determined for the subject though concealed for strategic reasons.

Determining WTP for privacy necessitates special consideration. Privacy is not a tangible good traded on designed or pre-configured markets. Rather it is a derived human right and treated differently with dependence upon culture and political regime. For instance, in Germany, privacy is often associated with the right to informational self-determination, i.e. the right of individuals to control, edit, manage, and delete their personal data and decide when, how, and to what extent it is communicated to others [7]. Hence privacy is more of a public good like environmental prevention or protection from crime provided by the administration though, like property rights, it is tightly linked to each individual in a society. The value individuals attribute to their personal data is a good indicator for their perceived importance of privacy. Since people do not trade personal data like they buy cars or pay for clothes but rather disclose personal data in exchange for using free (Internet) services WTP determination is somewhat different compared to classical goods or resources.

In the context of privacy research Willingness-to-Accept (WTA) payments for personal data and Willingness-to-Protect (WTPr) personal data are common variants of WTP. WTA observes the value someone would accept from a third party for making use of his personal data. The value someone would pay for data protection is measured by WTPr.

In the context of Online Social Networks, Krasnova et al. (2009) conducted a conjoint analysis to determine WTP of users being assured that the provider does not make use of their personal data [15]. The authors abstain from incentive compatibility and ask indirectly about stated privacy preferences of network users. From the background of their knowledge that the provider does not make use of personal data, users' average willingness to pay a monthly fee for continuing using the network was found to lie between €14.40 and €17.24.

Beresford et al. (2010) conducted a WTA study and showed that participants were willing to accept even small amounts of money for the disclosure of their data [16]. Participants were obliged to buy a product online and with disclosure of personal data they received a small discount. Receiving a discount for data disclosure determines WTP indirectly. Obligations to actually buy a product generate incentive compatibility.

Bauer et al. (2012) determined WTPr or the price someone would pay for data protection [17]. The authors conducted a study according to the BDM method [18]. Participants received information about a fictive scenario where all Facebook users were threatened by the loss of their profile and asked directly for their WTPr personal data. Incentive compatibility was implemented by means of a lottery drawing.

Acquisti et al. (2009) examined participants receiving shopping vouchers where the value of each voucher depended on personal data disclosure [19]. Participants were offered a price for their personal data. The more they were willing to disclose information the higher was the voucher value. According to this strategy participants were indirectly asked about their WTA payments for personal data and incentivized by the possibility to actually redeem vouchers. Interestingly, the authors were able to show that WTA payments for personal data is higher than WTPr personal data. Data protection and data sales are two different scenarios and with dependence upon contextual situation people value their data differently. Likewise, it appears reasonable that stated and revealed preferences are contingent upon observational context.

In summary, there are direct and indirect methods as well as incentive compatible and incentive incompatible means for determining privacy preferences. Direct methods are cognitively more demanding as personal data is not a tradable good. Participants may not have a clue about the monetary value of their data. Moreover, for measuring cognitive dissonance the monetary value is not of interest. What counts is the deviation between stated and revealed preferences. Obligations to buy products under consideration of discounts, vouchers, or other benefits offered in exchange for data disclosure create incentives to reveal true preferences. However, in most economic transactions and even more in Online Social Networks or other data sharing scenarios based on reciprocity, data disclosure is a side effect and there is no explicit trade-off between product and data. Incentive compatible methods create a highly artificial observational context and it is questionable if obligations trigger the

revelation of true preferences. From the assumption of preferences being constructed with the decision made incentive compatibility appears even more suspicious. Last but not least, some authors argued that auctions and lotteries bear understanding problems and, instead of avoiding strategic valuations, they cause strategic behavior [20].

4 Order Effects in Observations of Preferences

In many disciplines such as linguistics or cognitive psychology it is well known that the meaning of words or concepts depend on the context of their usage. Homonyms such as apple are good examples of context-dependent semantics. In the context of a eating a fruit apple has a different meaning compared to the context of a using a computing device. In his Philosophical Investigations Wittgenstein famously noticed that language has more than a denotative function. Speaking words or asking questions is always part of an activity, or form of life [21]. Likewise, determining WTP through surveying participants is a highly context-dependent activity. Here it was shown that the order of questions posed in a survey changes answering behavior of participants [22]. In other words, WTP is influenced by order effects which are expected to be a rule rather than an exception [23].

In several WTP surveys observations of proximal and distal order effects have been made [24]. Proximal order effects influence WTP according to the sequential measurement of successive questions. Distal order effects influence WTP due to information provided, or questions asked, outside of the WTP questions posed. For instance, asking participants about the price they would charge for disclosing their address details prior to asking them for the price they would charge for revealing their income level is supposed to reveal proximal order effects. WTP is expected to differ from the WTP where questions are ordered the other way round. If participants assume that income level will be associated with address details their valuation is expected to differ from the situation where they do not draw this association.

Distal order effects stem from information provided prior to WTP questions, or demographic questions determining participants' age and gender. For instance, values for WTA or WTPr of users in Online Social Networks are supposed to differ with respect to a scenario where the network provider is named, e.g. Facebook, or a scenario where the provider is kept incognito.

In the literature several attempts have been made to design surveys in a way that prevents order effects. Bateman et al. (2001) ask participants to value product bundles where each bundle is ranked according to its size [25]. Top-down rankings list large bundles on top, whereas bottom-up rankings show products descending from small to large bundles. The authors argued that WTP is higher for product bundles following up on smaller bundles. Vice versa, WTP is lower for bundles valued after the previous judgment of a larger bundle. A potential explanation of such proximal order effects is a successively changing visible choice set. A visible choice set is a complete list of goods participants have in mind during the whole surveying process. The direction of goods to be valued, i.e. top-down or bottom-up, but also changing number and type of goods, influences perceptions of visible choice sets. To prevent proximal order effects

occurring with dependence upon changing visible choice sets (stepwise disclosure), participants should be informed about all products to be valued prior to revealing their WTP (advanced disclosure).

As a mean to avoid order effects a priori advanced disclosure substantiates the role of knowledge about the good to be valued. Several studies witness a decreasing likelihood of order effects when participants are used to the particular good in question [26]. For instance, Boyle et al. (1993) found that WTP of people inexperienced in white-water rafting is prone to order effects while WTP of rafting professionals is not [27]. Kartmann et al. (1996) proved the absence of order effects in surveys where patients had to state WTP for treatments they were familiar with [28]. Transferred to the privacy paradox, familiarity with one's own privacy preferences is likely to have an impact on cognitive dissonance. If advanced disclosure indeed influences privacy preferences remains an empirical question.

Batemann et al. (2001) and Bateman et al. (2003) argued that in sequential answers to successive WTP questions order effects are likely to occur if participants state their willingness to pay for additional or extra features of previously judged goods [22,25]. Such inclusive lists present goods not as alternatives to be valued independently of each other, but as associated features where the value depends on previous judgments of related goods. Exclusive lists are countermeasures and present goods independently of each other. In other studies Kumar et al. (1991) and Chrzan (1994) conducted a conjoint analysis where product profiles are presented with varying attributes [29,30]. The authors showed that the order of product attributes influences WTP.

As discussed by Cai et al. (2011) distal order effects have an impact on survey results [24]. Information and descriptions provided prior to, or outside of, WTP questions is subject to varying interpretations. Therefore, the authors suggest randomizing all the information provided in the survey. Others found incentive compatible methods to be susceptive to distal order effects. Clark et al. (2008) argued that despite the purpose of the BDM method to guide participants in revealing their true WTP, often, people are misguided in understanding that stating their true WTP is in their own best interest [26]. Enlightening participants about what is best for them is a difficult exercise as people are quite unique in their ability to learn and apprehend the purpose behind inquiry. Teaching participants to act compatibly with the incentive provided is hard to control. Similar issues arise for auction mechanisms. Avoiding distal order effects in Vickrey auctions is difficult since explaining incentives behind second-price sealed-bid, i.e. exhausting WTP, also requires participants to learn and apprehend the purpose behind inquiry. Moreover, if participants interact among each other during auction experiments distal order effects are even harder to avoid.

5 A Design for Observations of Dissonant Privacy Preferences

Based on the discussion of WTP studies and order effects in the previous sections a design for measuring preferences in privacy decision making is developed in the

Table 1. Assignment of WTP Questions and Tasks to Treatment Groups

Treatment 1	Task 1	Question A Question B	Treatment 2	Task 1	Question B Question A
	Task 2	Question D Question C		Task 2	Question C Question D
Treatment 3	Task 2	Question C Question D	Treatment 4	Task 2	Question D Question C
	Task 1	Question B Question A		Task 1	Question A Question B

remainder[1]. Individuals experience dissonant states if data valuation (stated preferences) and disclosure behavior (revealed preferences) do not commute. Two basic methodological assumptions derive from the discussion so far. First, the design for measuring preferences abstains from naive realism, i.e. the claim that true preferences are out there readily determined. Second, the design assumes that observing order effects is a necessary condition for dissonant privacy decisions.

There are four questions A, B, C, and D. Questions A, C determine stated preferences for two similar scenarios. Likewise, questions B, D measure revealed preferences. A and B (scenario Facebook) measure the same type of data, i.e. name, address, and phone number, whereas C and D (scenario Amazon) also ask for the same type of data, i.e. bank account, monthly income, and credit card information.

All questions are assigned to two different tasks. Task 1 comprises questions A and B (scenario Facebook); task 2 is composed of questions C and D (scenario Amazon). Each task represents a sequential order of two successive questions. There are four treatment groups each of which is assigned two tasks where the order of the two questions within each task either starts with a stated preference question or a revealed preference question dependent upon the other tasks. For instance, participants in treatment group 1 start in task 1 with a stated preference question and in task 2 with a revealed preference question. The variation of questions and tasks is meant to counteract order effects. The reason for having four questions, or two scenarios (Facebook and Amazon), is that order effects can be determined for each individual of a treatment group. With only two questions, or one scenario (Facebook or Amazon), determining order effects would be possible solely between individuals of different groups. Individuals of one group would face only one order context, i.e. either AB or BA (cf. section 2).

Drawing from other studies observing privacy decision making, stated privacy preferences are best determined by surveying people about the degree to which they agree or disagree with a third party making use of their personal data (agreement). The degree to which people are concerned or agree with a third party constraining their right to privacy by making use of their data accommodates the fact that privacy

[1] A first version of the survey implementation is available at `https://www.elearning.uni-freiburg.de/ofb/iig/, password: abba`

is not a tradable good. Surveying participants indirectly appears reasonable since the real monetary value someone would charge a third party for making use of his personal data is not of interest.

Question A: I am concerned about the following data being used by Facebook for commercial purposes. [1: I disagree - 7: I strongly agree]
1. phone number
2. address details
3. name

Question C: I am concerned about the following data being used by Amazon for commercial purposes. [1: I disagree - 7: I strongly agree]
1. bank account
2. monthly income
3. credit card information

Revealed preferences are determined by measuring the degree to which someone complies with data disclosure as a necessary requirement for continuing using a service after terms and conditions have been changed (compliance). The degree of compliance mirrors the attractiveness of the service and the willingness to disclose personal data in exchange for using the service.

There are several reasons for not choosing an incentive compatible method to determine revealed preferences. As discussed in section 3 obligations to buy a product in line with discounts or vouchers received in exchange for data disclosure creates an artificial context. Moreover, as discussed in section 4, incentive compatible methods are prone to order effects due to their attempt to communicate the incentive behind inquiry. Since participants differ substantially in their ability to learn and comprehend the purpose of determining their true WTP, distal order effects would be introduced into the survey design. Apart from incentives, surveying participants indirectly appears also reasonable here. The real monetary value for continuing using the service is not of interest and variability is reduced.

Question B: Suppose Facebook changes its terms and conditions. To continue using the service all members are required to disclose their personal data truthfully. In case of false information the provider will close the respective account. Would you disclose the following data truthfully in order to continue using the service? [1: No - 7: Yes]
1. phone number
2. address details
3. name

Question D: Suppose Amazon changes its terms and conditions. To continue shopping online on Amazon members are required to disclose financial data. In case of false information the provider will close the respective account. Would you disclose the following data truthfully in order to continue using the service? [1: No - 7: Yes]
1. bank account
2. monthly income
3. credit card information

The following measures are applied to deal with order effects. First, participants accessing the survey are randomly assigned to one of the four treatment groups. According to the order of tasks as shown in Table 1 participants state and reveal their privacy preferences. Randomization of group assignments is meant to avoid order effects built into the design. Order effects can be determined for each individual by comparing WTPs for both scenarios, or they are observed on an aggregate level for each group. Additionally, WTP is compared among the four treatment groups. Second, before stating and revealing privacy preferences, participants receive information about the independence of each of the questions. Avoiding inclusive lists and semantic interdependence of questions explicitly counteracts order effects. Third, prior to WTP questions participants receive information about the process of the survey, i.e. information about the number of questions and what the questions are actually about. Advanced disclosure counteracts proximal order effects which are due to diverging visible choice sets. Finally, before the WTP questionnaire starts, participants' experience with the services is determined by asking for their familiarity with the subject. In the analysis phase familiarity as well as demographics will help clustering data according to their impact on order effects.

6 Conclusion and Future Work

Disclosure and dissemination of data has gained momentum. Incentives for using Internet services are tremendous and most Web users would not like to live without being online anymore. Nevertheless, most Internet users are deeply concerned about their digital future and the fate of their personal data. Being dissonant with regard to attitudes, opinions, and behaviors is omnipresent and not unusual for digital citizens. Here it was argued that cognitive dissonance accounts for the privacy paradox and that it can be measured in terms of order effects in sequential observations of successive privacy decisions. Existing WTP studies for determining the value of privacy have been discussed as well as design principles to avoid order effects built into the survey design. Eventually, a design artifact for measuring order effects in sequential privacy decisions has been proposed and implemented. Conducting the survey and presenting analysis results is planned for the near future.

The implications for follow-up studies are striking. Determining the degree of dissonance opens the arena for studies on the effectiveness of Transparency Enhancing Technology (TET). The guiding design principle for TET 2.0 to influence cognitive dissonance may become incentives, e.g. recommendations or search results, which do not fit putative preferences derived inferentially based upon individual and collective online behavior [31]. Large incentives keep cognitive dissonance low and thus may hinder change instead of reinforcing it. Accordingly, effectiveness of TET 2.0 hinges on the right magnitude of incentives imposed upon individual online behavior.

References

1. Acquisti, A., Grossklags, J.: Privacy and Rationality in Individual Decision Making. IEEE Security & Privacy 3, 26–33 (2005)
2. Acquisti, A., Grossklags, J.: Losses, Gains, and Hyperbolic Discounting: An Experimental Approach to Information Security Attitudes and Behavior. In: 2nd Annual Workshop on Economics and Information Security, WEIS, vol. 3 (2003)
3. Cusumano, M.A.: Platform wars come to social media. Commun. ACM 54, 31–33 (2011)
4. Müller, G., Flender, C., Peters, M.: Vertrauensinfrastruktur und Privatheit als ökonomische Fragestellung. In: Buchmann, J. (ed.) Internet Privacy - Eine multidisziplinäre Bestandsaufnahme/ A Multidisciplinary Analysis. Springer (2012)
5. Flender, C., Müller, G.: Type Indeterminacy in Privacy Decisions: The Privacy Paradox Revisited. In: Busemeyer, J.R., Dubois, F., Lambert-Mogiliansky, A., Melucci, M. (eds.) QI 2012. LNCS, vol. 7620, pp. 148–159. Springer, Heidelberg (2012)
6. Festinger, L.: A Theory of Cognitive Dissonance. Stanford University Press (1957)
7. Westin, A.: Privacy and Freedom. Atheneum, New York (1967)
8. Hevner, A., March, S., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly 28, 75–105 (2004)
9. Deuker, A.: Addressing the privacy paradox by expanded privacy awareness – the example of context-aware services. In: Bezzi, M., Duquenoy, P., Fischer-Hübner, S., Hansen, M., Zhang, G. (eds.) Privacy and Identity. IFIP AICT, vol. 320, pp. 275–283. Springer, Heidelberg (2010)
10. Spiekermann, S., Grossklags, J., Berendt, B.: E-privacy in 2nd generation e-commerce: privacy preferences versus actual behavior. In: Proceedings of the 3rd ACM Conference on Electronic Commerce, pp. 38–47. ACM (2001)
11. Awad, N., Krishnan, M.: The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS Quarterly, 13–28 (2006)
12. Norberg, P., Horne, D., Horne, D.: The privacy paradox: Personal information disclosure intentions versus behaviors. Journal of Consumer Affairs 41, 100–126 (2007)
13. Völckner, F.: An empirical comparison of methods for measuring consumers' willingness to pay. Marketing Letters 17(2), 137–149 (2006)
14. Miller, K., Hofstetter, R., Krohmer, H., Zhang, J.: How should consumers' willingness to pay be measured? An empirical comparison of state-of-the-art approaches. Journal of Marketing Research 48(1), 172–184 (2011)
15. Krasnova, H., Hildebrand, T., Guenther, O.: Investigating the value of privacy in online social networks: Conjoint analysis. In: Proceeding of ICIS 2009 (2009)
16. Beresford, A., Kübler, D., Preibusch, S.: Unwillingness to pay for privacy: A field experiment. Wissenschaftszentrum Berlin für Sozialforschung (2010)
17. Bauer, C., Korunovska, J., Spiekermann, S.: On the value of information – What Facebook users are willing to pay. In: Proceedings of the 20th European Conference on Information Systems, ECIS 2012 (2012)
18. Becker, G., DeGroot, M., Marschak, J.: Measuring utility by a single-response sequential method. Behavioural Sciences 9(3), 226–232 (1964)
19. Acquisti, A., John, L., Loewenstein, G.: What is privacy worth? In: 21st Workshop on Information Systems and Economics, WISE (2009)
20. Schreier, M., Werfer, J.: Auktionen versus Lotterien: Ein empirischer Vergleich zur Messung von Zahlungsbereitschaften. Die Betriebswirtschaft 67, 22–40 (2007)
21. Wittgenstein, L.: Philosophical Investigations (1953)

22. Bateman, I., Powe, N.: Ordering effects in nested 'top-down' and 'bottom-up' contingent valuation designs. Ecological Economics 45(2), 255–270 (2003)
23. Atmanpacher, H., Römer, H.: Order effects in sequential measurements of noncommuting psychological observables. Journal of Mathematical Psychology 56(4), 274–280 (2012)
24. Cai, B., Cameron, T., Gerdes, G.: Distal order effects in stated preference surveys. Ecological Economics 70(6), 1101–1108 (2011)
25. Bateman, I., Cooper, P., Georgiou, S., Poe, G.: Visible choice sets and scope sensitivity: an experiment and field test of study design effects upon contingent values. In: Proceedings of the 11th Annual Conference of the European Association of Environmental and Resource Economics (2001)
26. Clark, J., Friesen, L.: The causes of order effects in contingent valuation surveys: An experimental investigation. Journal of Environmental Economics and Management 56(2), 195–206 (2008)
27. Boyle, K., Welsh, M., Bishop, R.: The role of question order and respondent experience in contingent-valuation studies. Journal of Environmental Economics and Management 25(1), 80–99 (1993)
28. Kartman, B., Stalhammar, N., Johannesson, M.: Valuation of Health Changes with the Contingent Valuation Method: A Test of Scope and Question Order Effects. Health Economics 5(6), 531–541 (1996)
29. Kumar, V., Gaeth, G.: Attribute order and product familiarity effects in decision tasks using conjoint analysis. International Journal of Research in Marketing 8(2), 113–124 (1991)
30. Chrzan, K.: Three kinds of order effects in choice-based conjoint analysis. Marketing Letters 5(2), 165–172 (1994)
31. Celko, M.: Die Echokammer. Abstrakt: Die grosse Gemeinschaft. W.I.R.E. Collegium Helveticum 9, 25–31 (2012)

A User Privacy Protection Technique for Executing SQL over Encrypted Data in Database Outsourcing Service

T.B.P. Hue[1], D.N. Thuc[1], T.B.D. Thuy[1], Isao Echizen[2], and S. Wohlgemuth[2]

[1] University of Science, VNU - HCMC
227 Nguyen Van Cu, District 5, Ho Chi Minh City, Vietnam
{ptbhue,ndthuc,dtbthuy}@fit.hcmus.edu.vn
[2] National Institute of Informatics
2-1-1 Hitotsubashi, Chiyoda-ku, Tokyo, Japan
iechizen@nii.ac.jp, sven.wohlgemuth@me.com

Abstract. The fact that the data owners outsource their data to external service providers introduces many security and privacy issues. Among them, the most significant research questions relate to data confidentiality and user privacy. Encryption was regarded as a solution for data confidentiality. The privacy of a user is characterized by the query he poses to the server and its result. We explore the techniques to execute the SQL query over the encrypted data without revealing to the server any information about the query such as the query type or the query pattern, and its result. By implementing all the relational operators by using the unique selection operator on the server-side database with a constant number of elements in each time of selection, our proposal can defeat against the statistical attacks of the untrusted server compromising data confidentiality and user privacy. Experimental evaluation demonstrates that our proposal less affects the system's performance and is applicable in the real world.

Keywords: Database outsourcing, database encryption, user privacy, access pattern privacy, access privacy.

1 Introduction

Amount of data held by organizations is increasing quickly and it often contains sensitive information. Management and protection of such data are expensive. An emerging solution to this problem introduces a new paradigm called *database as a service* (DAS), in which the database of an organization is stored at an external service provider. The advantages of DAS are cost savings and service benefits. There are three main entities in the DAS scenario (Fig. 1): (1) *Data owner:* individual or organization that is the subject of the data made available for controlled external use (2) *User*: individual or organization that requests data from the server (3) *Server*: organization that receives the data sent from the data owners and makes it available for distribution to users.

In DAS scenario, however, sensitive data, which is now stored on a site that is not under the direct control of the data owner, can be put at risk. Moreover, the data request of user can be revealed to the untrusted server to violate the privacy of the

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 25–37, 2013.

user. Therefore, the data confidentiality and user privacy need to be taken into account. To ensure data confidentiality, the data owner needs to hide the database's content before outsourcing it to the service provider. We also know that the privacy of a user is characterized by the query he poses to the server and its result. It is necessary to protect both the query and its result from the unauthorized parties (such as untrusted server) to protect the user privacy.

Encryption was often considered as a solution for data confidentiality ([2], [4], [5], [6]). The order preserving encryption scheme supported the equality and range queries over the encrypted data [4]. Other work on privacy homomorphism illustrated techniques for performing arithmethic operations (+, -, x, /) on encrypted data ([5], [6]). Hacigümüs et al. [2] proposed storing, together with the encrypted database, additional indexing information. By using the created index, the server could execute the queries over the encrypted data. There were four steps to process a query (Fig. 2): (1) the query Q posed by a user was translated by the query processor at the client site to its server-side representation Q^S (2) Q^S was sent to the server and was executed over the encrypted database (3) the result (in encrypted form) was sent to the client; it was decrypted and filtered out those tuples not satisfying the query condition (4) the final result was sent to the user by the client. All of the above-mentioned work revealed to the untrusted server the query type of users, which was the useful information for the server to predict the query being requested. Using these query execution techniques, the clients transfer to the server the same query patterns when the users pose the same query requests. The untrusted server may perform statistical attacks and exploit these query patterns. By correlating known public information with the frequent query patterns, together with the query types, the server can infer the users' trend of information, or more critical the users' trend of sensitive information, which violates the privacy of users [3].

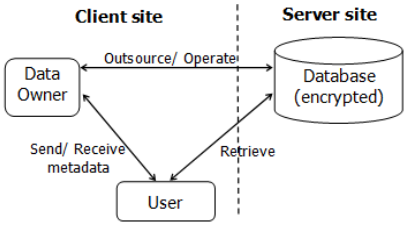


Fig. 1. Diagram of DAS

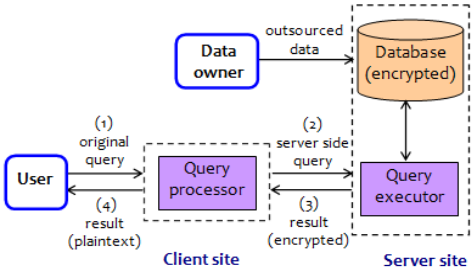


Fig. 2. Query execution process [2]

In this paper, we explore the techniques for executing SQL queries over the encrypted data without revealing to the server any information about the query or its result. Our proposal can defeat against the statistical attacks of the untrusted server compromising both data confidentiality and user privacy.

The remainder of this paper is organized as follows: section 2 presents the query execution solution for protecting the user privacy; section 3 presents the security analysis; section 4 is the experimental evaluation of our proposals; section 5 concludes the paper.

2 Our Proposed User Privacy Protection Technique

In this section, we propose a query execution technique which can protect the privacy of user. We adopt the database storage model and condition transformation technique proposed in [2].

2.1 Storage Model

For each relation r with the schema $R(A_1, A_2, \ldots, A_n)$, we store on the server an encrypted relation r^S with the schema $R^S(t^S, A_1^S, A_2^S, \ldots, A_n^S)$ where t^S stores an encrypted string that corresponds to a tuple in relation r, each A_i^S is a corresponding index to the attribute A_i that will be used for query processing at the server (Fig. 3). We can use any block cipher technique such as AES, RSA, Blowfish, etc., with the key size 128 bits. If there is unique user (also the data owner) in the system, we use one key for encrypting the whole database; otherwise, we use multiple keys which are managed by a key management mechanism [8, 9, 10]. The index is created based on the mapping function $Map_{R.Ai}(v)$, which will be defined as the following:

- The *partition function* which partitions the attribute's domain of values into disjointed partitions: $partition(r.A_i) = \{p_1, p_2, .., p_k\}$.
- The *identification function* $ident_{R.Ai}(p_j)$ which assigns an identifier to each partition p_j of attribute A_i.
- The *mapping function* which maps a value v in the domain of attribute A_i to the identifier of the partition to which v belongs: $Map_{R.Ai}(v) = ident_{R.Ai}(p_j)$, where p_j is the partition that contains v.

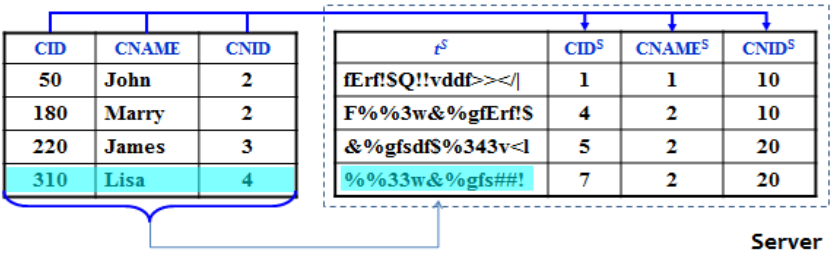


Fig. 3. Data storage model [2]

We use the operator D that maps the encrypted representation to its corresponding unencrypted representation. That is, $D(r^S) = r$. For differentiating the execution of an operation on the client site or on the server site, we denote the "S" in superscript form beside the operator with the suggestion to execute this operator on the server site. For example σ^S denote the selection operator is executed (on the encrypted data) at the server site. We denote R^+ for the set of all the attribute of r: $R^+ = \{A_1, \ldots, A_n\}$.

In our proposal, the principles of all the relational operators on each database management systems are unchanged. Besides the notation of the normal relational operators, such as $\prod$, σ, $\bowtie$, $\cup$, $\cap$, -, $\leftarrow$ which stand for the projection, the selection, the join, the union, the intersection, the set difference, the assignment respectively as

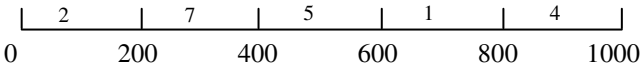
defined in [1], we use the additional ones: τ denotes for the sorting operator, and γ denotes for the grouping and aggregation operator.

2.2 Condition Transformation

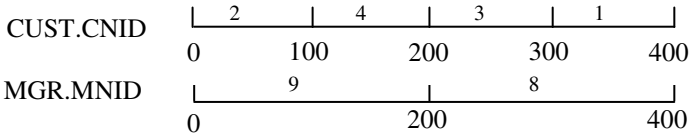
The condition mapping function, Map_{cond}, translates a condition from a user's query to its corresponding one over the server-side representation [2]. This section we only refer to random mapping of two popular types of condition, which will be used in the next section. For the more consideration, please refer to [2]. We will use these two relations for illustrations:

CUST (CID, CNAME, CNID)
MGR (MID, MNAME, MNID)

Attribute = Value: $Map_{cond}(A_i = v) \Rightarrow A_i^S = Map_{Ai}(v)$. For example, $Map_{cond}(CID = 250) \Rightarrow CID^S = 7$.



Attribute$_1$ = Attribute$_2$: $Map_{cond}(A_i = A_j) \Rightarrow \bigvee_{\varphi} (A_i^S = ident_{Ai}(p_k)$ và $A_j^S = ident_{Aj}(p_l))$, where φ is $p_k \in$ partition (A_i), $p_l \in$ partition(A_j), $p_k \cap p_l \neq \varnothing$.



For example, $Map_{cond}(CUST.CNID = MGR.MNID) \Rightarrow (CUST^S.CNID^S = 2 \wedge MGR^S.MNID^S = 9) \vee (CUST^S.CNID^S = 4 \wedge MGR^S.MNID^S = 9) \vee (CUST^S.CNID^S = 3 \wedge MGR^S.MNID^S = 8) \vee (CUST^S.CNID^S = 1 \wedge MGR^S.MNID^S = 8)$.

2.3 Solution for Protecting Access Pattern Privacy

Principles. Our solution was based on three principles: (1) all the relational operators (from the client query) are implemented by doing only the selection operator over the server-side database (2) we select (n + m) elements in each time of selection over the server-side database, where n and m are the parameters which determine the security level of our proposed system (3) minimizing the work done at the client side.

Conforming to the principle 1, the untrusted server cannot recognize the type of the query that is being requested. Respond to whatever the query type required by the user, the server simply does the selection. The algorithm Select_NTimes is used by the client for dispatching the selection request(s) to the server and receiving the result in encrypted form. The number of elements requested in each time depends on the total number of index values being requested and the values of n and m.

Principle 2 prevents the server from doing the statistical attacks to learn the frequent query pattern. In the case the users request the same query, the sets of

elements the corresponding client request from the server in each time of selection are different with the high probability. The appropriate values of n and m should be suggested by the data owner or an expert in the field involved. We will analyze the security of our proposed techniques in section 3. Principle 3 keeps the spirit of database outsourcing service in which most of work should be done by the server.

Solution for Selecting Data from the Server. All the relational operations over the encrypted database will be implemented by using the Select_NTimes algorithm. Note that GetRand(n, I) is the function for getting randomly n elements from the set I while card(I) is the function for returning the cardinality of the set I.

Algorithm Select_NTimes(r(R), A, I, n, m) For selecting from the encrypted relation $r^S(R^S)$ of relation r(R) the tuples with the value at the attribute A^S belonging to the set of values I, $I \subseteq Ident(R.A)$

```
T = ident(r.A) – I; R = ∅
While card(I) > 0
Begin
        N = ∅
        If card(I) <= n then
            Begin
                L = I
                If card(I) < n then
                        N = GetRand(m + n – card(I), T)
                Else if card(I) = n then
                        N = GetRand(m, T)
                I = ∅
            End
        Else // card(I) > n
             Begin
                If card(I) <= n+m then
                Begin
                        L = GetRand(n, I);
                        N = GetRand(m, T)
                End
                Else // card(I) > n+m
                        L = GetRand(n+m, I)
                        I = I – L
             End // card(I) > n
        Z = L ∪ N
        R1 ← σ^S_{A^S ∈ Z}(r^S)
        R2 ← σ^S_{A^S ∈ L}(R1)
        R ← R ∪ R2
End While
Return R
```

The Select_NTimes algorithm operates in the following manner. Let T be the set of values of A^S except the values in I. If the cardinality of the requested set I is less than n, the client adds to I the values in T in order to have a set Z having the cardinality (n+m) in each time of selection. In the case the cardinality of I is equal to n, the client conforms to principle 1 to add m values in T more. In the case the cardinality of I is greater than n but less than (n+m), the client to get randomly n values from I and add together with m ones getting randomly in T for each time of selection over r^S. If the cardinality of I is greater than (n + m), the client flexibly selects (n +m) values randomly from I for each time of selection over r^S. When receiving the result returned from the server, the client should remove the spurious tuples for saving the cost for decrypting them (see Fig. 4).

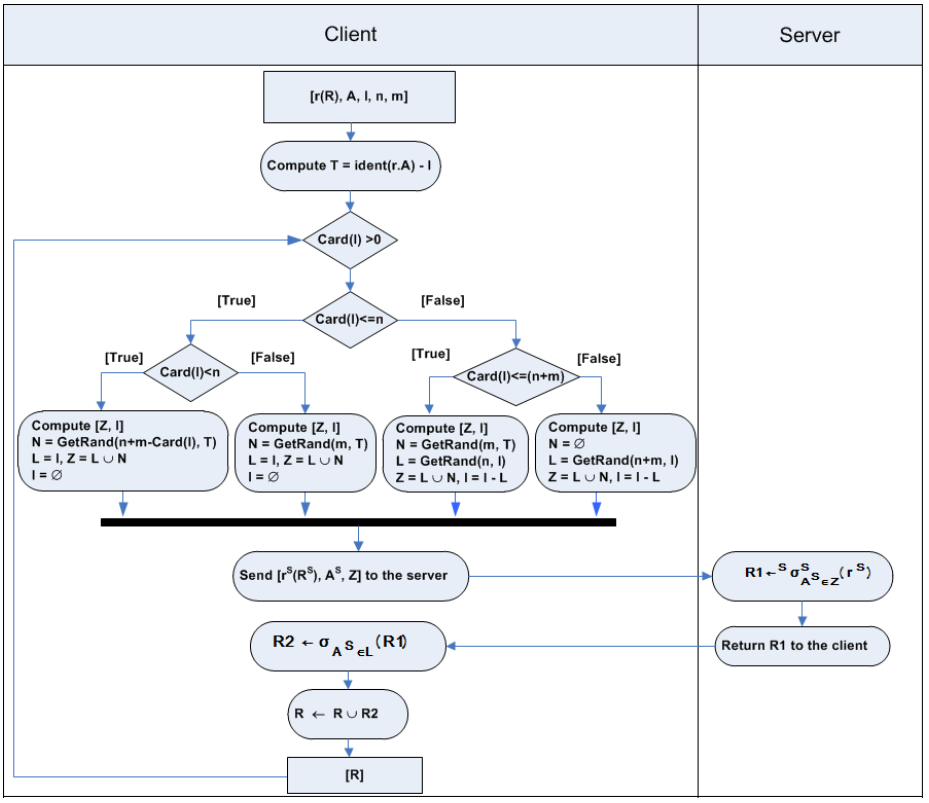


Fig. 4. Select_NTimes algorithm

We suggest decrypting the result one time after selecting from the server all the satisfied rows. By this way, we save the time of using resource of the client. However, it requires the client to store the result before decrypting all of them. An alternative way is to apply the client side operation to the tuples arriving over the answer stream as soon as they arrive without the need to store them.

We knew that sorting the input data is necessary for implementing operations such as join, union, intersection, duplicate elimination. Every tuple belonging to a single group of γ_L will be in a single group of $\gamma^S_{L'}$ computed by the server. The client only needs to consider tuples in a single group of $\gamma^S_{L'}$ when computing the groups corresponding to γ_L. The sort operator τ_L on the tuples having been grouped by the server can also be implemented efficiently using the merge-sort algorithm. These are the reasons why we design the algorithm, named Select_NTimes_Grouped, which has the same function as Select_NTimes except that the result set is grouped according to the specified attributes after every selection done by the server. Algorithm Select_NTimes_Grouped(r(R), A, I, n, m, L) is used for selecting from encrypted relation $r^S(R^S)$ corresponding to the relation r(R) the tuple(s) having the value(s) at attribute A^S belonging to the set of values I. The returned result is grouped by L^S, the correspondence of L.

By using Select_NTimes_Grouped algorithm, the operations that need the input data to be grouped or sorted (such as join, grouping and aggregation, sort, duplicate elimination) be implemented efficiently.

Implementation of Relational Operators

- **Selection operator**

Algorithm Selection (r(R), A, C, n, m) For selecting from relation r(R) the tuple(s) having the value(s) at attribute A satisfying condition C

$Map_{cond}(C) \Rightarrow A^S \in I$;
result ← Select_NTimes(r(R), A, I, n, m)
return σ_C(D(result))

We explain the above implementation using the example: $\sigma_{CID = 500}$ (CUST). $Map_{cond}(CID = 500) \Rightarrow CID^S \in \{5\}$, which means I = {5}. Suppose that the parameters for attribute CID of relation CUST are n = 2 and m = 2, according to the algorithm Select_NTimes, because card(I) = 1 < n (n = 2), the client will choose (m+n – card (I)) = (2+2 -1) = 3 elements randomly from ident(CID) – I = {2, 7, 5, 1, 4} – {5} = {2, 7, 1, 4}. Suppose that 2, 1 and 4 are selected, which means N = {2, 1, 4}. The client requests the set L ∪ N = {5, 2, 1, 4} rather than requests only I = {5}. The client then decrypts the result and filters out the tuples satisfying the condition C.

- **Join operator**

Algorithm Join (r(R), t(T), C, n, m) For returning the result of $r \bowtie_C t$, C is a θ- join condition
I, J contain all possible partitions of A_i, A_j that exists at least one pair of them may provide some values of A_i and A_j that can satisfy the condition C: $A_i \theta A_j$.
result1 = Select_NTimes_Grouped(r(R), A_i, I, n, m, A_i)
result2 = Select_NTimes_Grouped(t(T), A_j, J, n, m, A_j)
result = σ_C(D(result1 $\bowtie_{Map_{cond}(C)}$ result2))

return result

For instance, with the join condition C: CUST.CNID = MGR.MNID, Map_{cond}(CUST.CNID = MGR.MNID) $\Rightarrow$ (CUST^S.CNID^S = 2 $\wedge$ MGR^S.MNID^S = 9) $\vee$ (CUST^S.CNID^S = 4 $\wedge$ MGR^S.MNID^S = 9) $\vee$ (CUST^S.CNID^S = 3 $\wedge$ MGR^S.MNID^S = 8) $\vee$ (CUST^S.CNID^S = 1 $\wedge$ MGR^S.MNID^S = 8)

For using our proposed JOIN algorithm, we identify that A_i is CNID, A_j is MNID, I = {2, 4, 3, 1}, J = {9, 8}. The client selects all the rows from the relation r^S satisfying the condition 'CNID^S in I', which resulted in result1. The client also selects all the rows from the relation t^S satisfying the condition 'MNID^S in J', which resulted in result2. The client executes the join operation between result1 and result2 with the join condition is Map_{cond}(C), which resulted in result. The client continues executing the decryption operator on result and selecting the rows satisfying the condition C.

- **Projection operator**

Algorithm Projection (r(R), n, m, L) For returning the projection of r(R) on the projection attribute(s) L

$A \in R^+$; I = ident(A)
result = Select_NTimes(r(R), A, I, n, m)
return $\prod_L$(D(result))

The projection operator cannot be implemented on the server because each tuple of r is encrypted together into a single string in the t^S attribute of r^S. After selecting all the rows of the relation r, the client decrypts the result and performs the projection.

- **Grouping and aggregation operator**

Algorithm Group_Aggregation(r(R), L, n, m) For returning the values of aggregation functions operating on each group
$L = L_G \cup L_A$, L_G contains attributes on which the grouping is performed; L_A corresponds to a set of aggregation operations.
I = ident(A_i); $A_i \in L_G$
result = Select_NTimes_Grouped(r(R), A_i, I, n, m, L_G)
return γ_L(D(result))

The grouping and aggregation operation is denoted by $\gamma_L(r)$ where $L = L_G \cup L_A$. L_G is the list of attributes on which the grouping is performed while L_A is the set of aggregation operations. The server does not perform any aggregation corresponding to L_A. It returns all the rows of the relation r in responding to the client's request using Select_NTimes_Grouped algorithm. This result has been grouped by the server on the corresponding group of L_G. The client decrypts this result and performs the grouping operation and computing the aggregation functions specified in L_A.

- **Sort operator**

Algorithm Sort (r(R), L, n, m) For sorting the tuples of r(R) by L
I = ident(A_i); $A_i \in R^+$
result = Select_NTimes_Grouped(r, A_i, I, n, m, L)
return τ_L(D(result))

The sorting operator is implemented similarly to the grouping operator. The client firstly selects all the rows of the relation r using the Select_NTimes_Grouped algorithm. This result has been grouped by the server on the encrypted attributes of those in L. The client then decrypts the result, performs the sorting operation on the attributes in L. If the mapping functions of the attributes in R^+ are all order-preserving, the grouping operation operated on each part of the result returned by the Select_NTimes_Grouped should be replaced by a corresponding sorting operation for saving the cost at the client. The reason is that the result returned by the server is presorted within the partition. Sorting the result is a simple local operation over a single partition.

The following three set operators must be executed on the two compatible relations. These operators are implemented by the same manner. They cannot be executed by the server because on the encrypted form of the relations r and t, it is impossible to tell whether or not a given tuple satisfies the current operator. The client firstly selects all the rows of two relations r and t using the Select_NTimes_Grouped algorithm, except the union operator (without duplicate elimination) using the Select_NTimes algorithm. The client then decrypts the results and performs the corresponding operation.

- **Set operators**

Algorithm Difference (r(R), t(T), n, m) For returning the difference between r(R) and t(T)
$A \in R^+$; $B \in T^+$; A and B have the same domain value
I = ident(A); J = ident(B)
result1 = Select_NTimes_Grouped(r(R), A, I, n, m, A)
result2 = Select_NTimes_Grouped(t(T), B, J, n, m, B)
return D(result1) – D(result2)

Algorithm Algorithm Union (r(R), t(T), n, m) For returning the union of r(R) and t(T)
$A \in R^+$; $B \in T^+$; A and B have the same domain value.
I = ident(A); J = ident(B)
result1 = Select_NTimes(r(R), A, I, n, m, A)
result2 = Select_NTimes(t(T), B, J, n, m, B)
return D(result1) $\cup$ D(result2)

Algorithm Intersect (r(R), t(T), n, m) For returning the intersection between r(R) and t(R)
$A \in R^+$; $B \in T^+$; A and B have the same domain value.
I = ident(A); J = ident(B)
result1 = Select_NTimes_Grouped(r(R), A, I, n, m, A)
result2 = Select_NTimes_Grouped(t(T), B, J, n, m, B)
return D(result1) $\cap$ D(result2)

3 Security Analysis

Our proposed system use two parameters n and m. n is used for preventing the server from predicting the query type of the user based on the number of values to be requested. For example, when requesting an exact match query, the number of requested values is 1, which contrasts to the projection query with a larger number of requested values. m is used with the purpose of making noise to prevent the server from predicting the query type in the case the cardinality of the requested set is small (card(I) < n+m).

There is the trade-off between the security level and the communication and computation cost in our proposed system. The higher the value m is, the harder for the server to predict exactly the query is being executed. When the value of m is large enough, the probabilities of being selected of all the values in the considering domain are similar to each other, which creates difficulty for the server to predict the query type or the query pattern of users. However, the high value of m affects the performance of the system.

In the case the cardinality of I is greater than (m+n), each time of selection the client choose one set in N sets of values for selecting from server: $N = C_{card(I)}^{n+m}$. For preventing the case the server finding the intersection of the requested sets for predicting the query pattern, the value of N must be large enough. The higher the value of N, the more secure the system is. N becomes maximum if (n+m) approximates to card(I)/2.

Every attribute to which there may be have the query relates to should be set the values of n and m. For security reason, all the selection conditions done on an attribute should use the same values of n and m.

4 Experimental Evaluation

We present the experimental evaluation of our proposal. We implemented our proposed query execution method and the one suggested by Hacigümüs et al. [2], called Hacigümüs, and compared the query execution time between them.

By utilizing TPC-H benchmark [7], we generated two relations containing information about customers and mangers: CUST (CID, CNAME, CNID) and MGR (MID, MNAME, MNID). These attributes mean customer's identity, customer's name and customer's nation identity. The attributes of MGR relation have the same meanings as ones in CUST. We generated 150000 rows for the relation CUST, with the CID ranged from 1 to 150000. We generated 1000 rows for the relation MGR. The nation identity attributes (CNID and MNID) ranged from 1 to 25.

Our experiments were carried out on an Intel© Core2 Duo Processor P8700 2.53GHz, 4GB RAM. Relevant software components are Windows 7 as the operating system, SQL Server 2005 as the database management system and Microsoft Visual Studio C++ 2008 as the programming language. We used the equi-width technique to partition the domain of attributes CID, CNID, MID and MIND. The domains of

attributes CID and MID were partitioned into fragments, each fragment contained 49 integer values. The domain of attribute CNID was partitioned into 5 fragments while the domain of attribute MNID was partitioned into 3 fragments. We considered four queries: one exact match selection, one range selection, one join and one projection.

Q1: SELECT * FROM CUST WHERE CID = 500; Q2: SELECT * FROM CUST WHERE CID >= 500; Q3: SELECT * FROM CUST, MGR WHERE CNID = MNID; Q4: SELECT CID, CNAME FROM CUST;

For the query Q1, the condition after mapping was $CID^S = 10$. Firstly, we executed it 5 times using the execution process proposed by Hacigümüs et al. [2] and recorded the execution time. Secondly, by using our proposed access pattern privacy protection techniques, we executed Q1 with the value of n was 2 and the value of m run from 2 to 6, and computed the average execution time. We also repeated it 5 times. Thirdly,

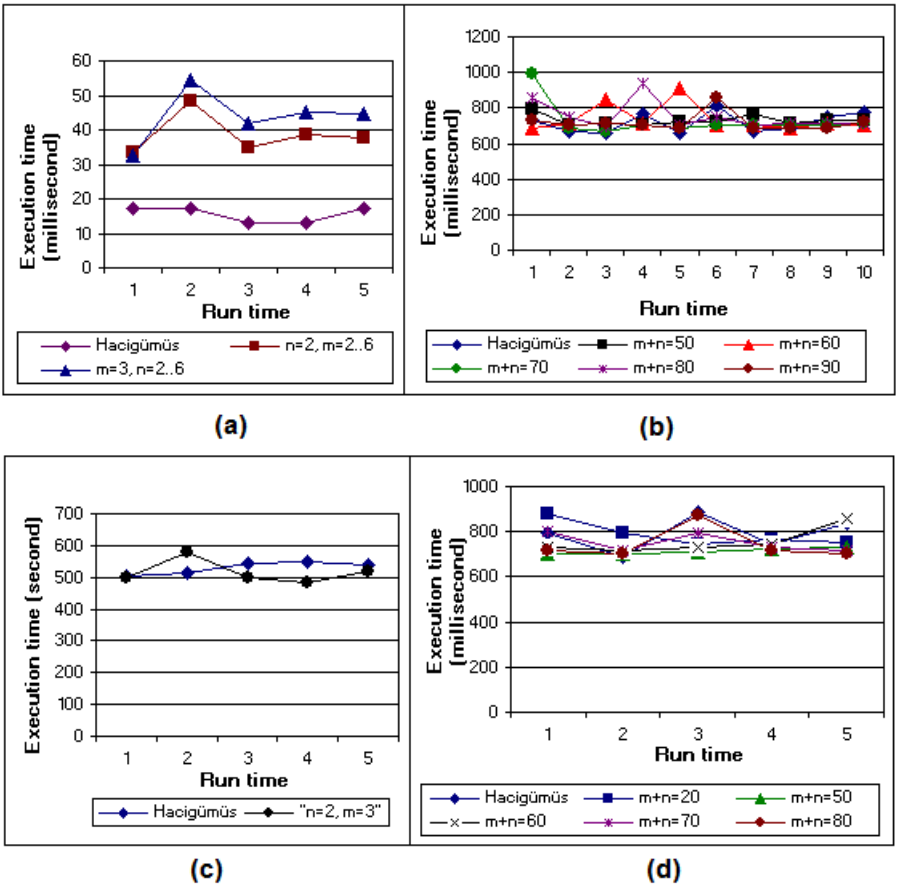


Fig. 5. Execution time when using our access pattern protecting solution comparing with that when using techniques of Hacigümüs el al.: (a) Q1 (b) Q2 (c) Q3 (d) Q4

we did the same things with Q1 as the second execution time with the value of m was 3 and n run from 2 to 6. For each pair of n and m, we run the query 10 times with different random sets of selection values according to the Select_NTimes algorithm. Fig. 5 (a) demonstrates that there is the small difference between the execution time of the two latter running times. The amount of execution time of these two running times is certainly higher than that of Hacigümüs et al. because the higher number of selection values were selected from the server.

For the query Q2, the condition after mapping was the set containing 290 values. We executed Q2 using Select_NTimes algorithm with the cardinality of the selection sets of values in each time of execution the selection (that was m+n) varied by 50, 60, 70, 80, 90. For each value of (m+n), we run the query 10 times with different random sets of selection values. Fig. 5 (b) demonstrates that the execution time when using our proposed execution technique is the same as that when using Hacigümüs' one. The differences in the execution time between the values of (m+n) are low during 10 times of running the experiment.

Executing Q3 by using Select_NTimes_Grouped algorithm (with n= 2 and m = 3) costs the same amount of time as that by using Hacigümüs' one, Fig. 5 (c). The result of the join operator contains 599988 rows which need more than 8 minutes to produce.

The result of the query Q4 contains all the rows of the CUST relation (150000 rows). We executed Q4 using Select_NTimes algorithm with the cardinality of the selection sets of values in each time of execution the selection (that was m+n) varied by 20, 50, 60, 70, 80. For each value of (m+n), we also run the query 10 times with different random sets of selection values. What we see on the Fig. 5 (d) is that the differences in execution time between the values of (m+n) are small, and the execution time of our proposed techniques is the same as that of Hacigümüs' one.

5 Conclusion

In this paper, we analyze the existing solutions for protecting data confidentiality and user privacy in DAS. The recent and well-known proposal of Hacigümüs el al. [2] is expressive but cannot defeat against the statistical attacks of the untrusted server, which may violate the data confidentiality and the user privacy. We propose the simple but robust technique for executing the relational operators over the encrypted database which can protect both the data confidentiality and the user privacy. Experimental evaluation demonstrates that our proposal less affects the system's performance and is applicable in the real world.

References

1. Elmasri, R., Navathe, S.B.: Fundamentals of Database Systems, 4th edn. Addison-Wesley (2004) ISBN 0-321-12226-7
2. Hacigümüs, H., Iyer, B.R., Li, C., Mehrotra, S.: Executing SQL over encrypted data in the database-service-provider model. In: SIGMOD, pp. 216–227 (2002)

3. Sion, R.: Towards Secure Data Outsourcing. In: Handbook of Database Security, pp. 137–161 (2008)
4. Agrawal, R., Kierman, J., Srikant, R., Xu, Y.: Order preserving encryption for numeric data. In: Proc. of ACM SIGMOD 2004, France (2004)
5. Boyens, C., Gunter, O.: Using online services in untrusted environments – a privacy-preserving architecture. In: Proc. of the 11th European Conference on Information Systems (ECIS 2003), Italy (2003)
6. Li, F., Hadjieleftheriou, M., Kollios, G., Reyzin, L.: Authenticated Index Structures for Aggregation queries in Outsourced Databases, Technical Report BUCS-TR-2006-011 (2006)
7. TPC-H. Benchmark Specification, `http://www.tpc.org`
8. De Capitani di Vimercati, S., Foresti, S., Jajodia, S., Paraboschi, S., Samarati, P.: Over-encryption: Management of Access Control Evolution on Outsourced Data. In: VLDB, pp. 123–134 (2007)
9. Atallah, M.J., Frikken, K.B., Blanton, M.: Dynamic and efficient key management for access hierarchies. In: Proc. of the ACM CCS, Alexandria, VA, USA (2005)
10. Hue, T.B.P., Luyen, G.N., Kha, N.D., Wohlgemuth, S., Echizen, I., Thuc, D.N., Thuy, T.B.D.: An Efficient Fine-grained Access Control Mechanism for Database Outsourcing Service. In: Proc. of the Int. Conf. on Information Security and Intelligent Control (ISIC 2012), pp. 67–71. IEEE Computer Society Press, Taiwan (2012) ISBN 978-4673-2586-8

Assessing Emotions Related to Privacy and Trust in Personalized Services

Ilias O. Pappas[1], Michail N. Giannakos[2], Panos E. Kourouthanassis[1], and Vassilios Chrissikopoulos[1]

[1] Department of Informatics, Ionian University, Corfu, Greece
{ilpappas,pkour,vchris}@ionio.gr
[2] Department of Computer and Information Science, Norwegian University of Science and Technology (NTNU)
michail.giannakos@idi.ntnu.no

Abstract. This study explores the dynamics of personalized services in online shopping, with regard to emotions, privacy and trust. The basic emotions of happiness and anxiety were chosen. A sample of 182 online shoppers was used to assess the effect of privacy and trust on their emotions through personalized services, and how these emotions ultimately affect their purchase intentions. The findings indicate that privacy affects anxiety while trust affects happiness, while both emotions have significant influence on customers' intention to buy through personalized services. The study concludes with theoretical and practical implications, limitations, and future research directions.

Keywords: Personalization, emotions, privacy, trust, anxiety, happiness, intention.

1 Introduction

Advances in technology give the opportunity to online vendors to offer high level of personalization, which can be used to communicate with customers in different levels. The collection and use of transactional, demographic and behavioral data makes it a great way to offer personalized services to every customer. Online vendors may offer a friendlier and individualized shopping experience, which eventually might increase customers' loyalty [1]. Online personalization has not been extensively addressed regarding customers' behavior [2].

Personalized services are based on customers' personal and private information, so if a customer wants to use such services he is obligated to share them. However, a level of trust that the service provider will behave ethically should exist. At the same time privacy concerns should be accounted for since they start affecting customers almost as soon as they share their data, reducing their purchase intentions [3]. The collection and use of private information have caused serious concerns about privacy invasion by customers [4] Previous studies have found that online shopping behavior is affected from customers' emotions [5]. Emotions, as a predictor of Information

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 38–49, 2013.

Technology use [6], may be a result of using personalized services. Although positive and negative emotions have been found to affect post purchase intentions [7], their role in personalized services still remains understudied.

This study provides a first insight into what factors affect the provision of personalized services in e-commerce environments. Previous studies have identified the importance of personalization in online shopping [8]. Likewise, scholars have shed light on the factors that affect adoption behavior of personalization [9]. Nevertheless, the role of emotions on this adoption behavior remains largely understudied. Drawn on the above, the objectives of this study are two-fold. On the one hand, we seek to investigate how happiness and anxiety affect adoption behavior of personalized services. On the other hand, we explore how privacy and trust issues shape the formulation of emotions on personalized environments. Privacy and trust issues are considered in an attempt to explain their relation with happiness and anxiety, and how these emotions affect customers' intentions. To this end, an empirical model is proposed and tested using structural equations modeling (SEM).

This paper is organized as follows. In the next section we review the existing literature on privacy, trust, happiness, anxiety and intention to purchase. In section 3, we present the theoretical foundation of the research model. In section 4, we present the methodology and the measures adopted for collecting data on the online shopping behavior. Section 5 presents the empirical results derived. In the last section of the paper, we discuss the findings and conclude by providing theoretical and practical implications and make recommendations for future research.

2 Literature Review

Online personalization refers to providing customers with tailored content and services based on knowledge obtained through service and user interactions [10]. In this study we employ the definition of Roberts [11] who defines personalization as "the process of preparing an individualized communication for a specific person based on stated or implied preferences" (p.462). Previous studies have identified the importance of personalization in online shopping and its effect on customers' behavior [2][12]. In order for personalization to be successful and to achieve its main goal, which is satisfied customers, different factors need to be taken into account. A vast overview on the subject is provided from Adoplhs and Winkelman [9], including user centric aspects, implementation and theoretical foundations. Nevertheless, emotions are not included.

In the ever growing field of online shopping customers' emotional reactions are common. Hedonic motivations have been found to affect customers' shopping online experience and their future intentions [13]. The different emotions that arise from online shopping can be affected or triggered by using personalized services. However, there is limited research on the different emotional aspects that occur from online shopping [14]. It has been argued that emotions are constituted of different constructs, although it is generally agreed that at least four are the basic emotions, namely

happiness, sadness, anxiety and anger [15]. Our study adopts the emotional constructs from Kay and Loverock [15]. Specifically we examine happiness and anxiety. Happiness is defined as the extent to which a person feels satisfied, excited and curious. Anxiety refers to the extent to which a person feels anxious, helpless, nervous and insecure.

People are expected to avoid behaviors that create anxious feelings and prefer those that give them happiness. In the more general area of IT use, happiness has been found to affect IT use positively, but a negative effect was found on task adaptation, which refers to how the user modifies they way something is done based on the technology used [6]. In other words, users are less happy when they have to change their personal preferences on how they complete a task in order to gain more benefits. Regarding anxiety, it was found a direct negative effect on IT use, and an indirect positive effect on IT use through social support [6]. This means that when users as for help from people they personally know (i.e. family, friends), their anxiety has a lower effect on IT use. It can be inferred that when the service is offered personally to a user, while based on individual preferences and tailored to the his needs, his anxiety will be reduced.

Previous studies point out the importance of anxiety while using computers [3]. Specifically, low levels of anxiety lead to more positive attitudes towards information sharing, essential for personalization. Moreover, anxiety has a negative effect on customers' intention to use mobile shopping [16]. In the context of computer learning, previous studies have found that higher levels of happiness and lower levels of anxiety may lead to increased computer use [15]. Anxiety has been studied extensively, focusing on system or computer anxiety [17]. However, anxiety in the context of personalized services is understudied.

Anxiety is very interesting to study as it provides insight into customers' general concerns regarding privacy [3]. Moreover, in order to develop long-term relationships with their customers, it is important for e-retailers to both develop and nurture consumer trust [18]. Both privacy and trust have been examined in the general context on online shopping, however their effects on basic emotions such as happiness and anxiety are understudied. Additionally, previous studies have examined the effect of privacy on personalization and information sharing. Collecting and using private information for personalization purposes has increased the privacy concerns of the customers [4]. It is proposed that if customers are given control over the use of their data, they are willing to share them because they feel their privacy will not be violated [3]. However, Brandimarte et al (2012)[19] found that offering users high control of their private data does not always lead to high privacy protection, because the sense of security that is created leads them to share even more information to a wider audience. The effects of privacy and trust on emotions need to be studied when customers use personalized services, because they might change depending on the offered services; how the services are offered, what they include and what the customer has to gain from them [20-22].

3 Hypotheses and Model Development

The aim of this study is twofold. First, we investigate the effects of privacy and trust on both happiness and anxiety. Next, we assess the effects of happiness and anxiety on intention to purchase while using personalized services.

3.1 Do Privacy Issues Shape Our Emotions?

Privacy is important for a customer that wants to create a relationship with an online vendor. Taking into account that personalized services are based on customers' personal data, privacy concerns become even more important in this relation. Lee and Cranage [23], examined the personalization-privacy paradox, where the better services a customer wants the more personal information he has to share, and found that high privacy issues increase customers' unwillingness to share such information and reduce their future intentions. On the other hand, customers are more willing to reveal personal data when they feel they can control their future use [3]. Pappas et al. [20] posit that high privacy concerns towards personalized services reduce customers' enjoyment, while Xu et al. [21] found that using personalized services might help customers to override their privacy concerns. Nonetheless, privacy issues towards online shopping, affect anxiety, which is likely to reduce customers' positive feelings [17]. Hence, we propose that:

H1: *Privacy will have a negative effect on happiness.*
H2: *Privacy will have a positive effect on anxiety.*

3.2 Do Trust Issues Shape Our Emotions?

Trust is critical for an online vendor to be successful, especially when personalized services are offered. Hwang and Kim [17] found that customers' affective reactions are related with trust. When referring to trust emotions are present [24]. Previous studies have showed that depending on the customers' involvement with an online vendor, factors such as satisfaction are decisive when fostering trust [25]. Taking into account that in order for personalized services to work, customers' involvement is needed and that satisfaction is closely related with emotions we propose that:

H3: *Trust will have a positive effect on happiness.*
H4: *Trust will have a negative effect on anxiety.*

3.3 How Emotions Influence Our Purchase Intentions?

Different emotions arise during consumption and affect customers' behavior. These emotions might either be positive or negative. Previous studies have showed that emotions' effect on intention might either be positive or negative [5][26]. Koo and Ju [27] found that pleasure and arousal that derive from atmospherics affect positively online shopping intention. Moreover, in the context of mobile shopping services anxiety was found to affect negatively behavioral intentions [15]. Besides, the more generalized positive and negative emotions, it is essential to examine the specific types-categories of emotions and how they affect customers' intentions while using personalized services. Consequently, we propose that:

H5: *Happiness will have a positive effect on intention to purchase.*
H6: *Anxiety will have a negative effect on intention to purchase.*

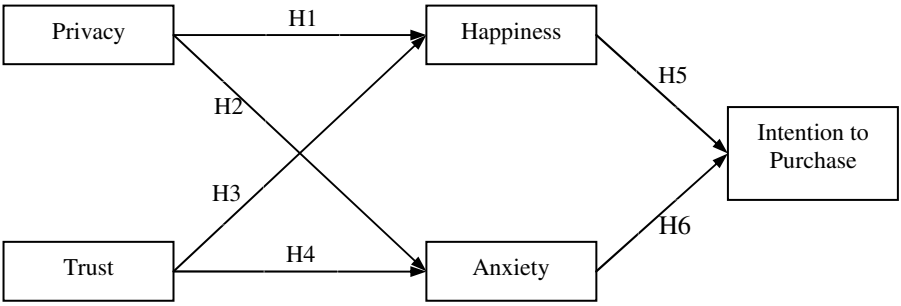


Fig. 1. Research Model

4 Methodology

4.1 Sample

Our research methodology included a survey conducted through the delivery and collection of individual questionnaires. It was made clear that there was no reward for the respondents and the participation was voluntary. The survey was executed in June-July 2012. We aimed at 600 (Greek) users of online shopping, 182 of which finally responded.

Table 1. Users' demographic profile

Demographic Profile		No	%
Gender	Male	98	53.8%
	Female	84	46.2%
Marital Status	Single	132	72.5%
	Married	45	24.7%
	Divorced	5	2.7%
Age	0-24	52	28.6%
	25-29	56	30.8%
	30-39	44	24.2%
	40+	30	16.5%
Education	Middle School	2	1.1%
	High School	22	12.1%
	University	78	42.9%
	Post Graduate	80	44%

As Table 1 shows, the sample of respondents was composed of almost equally men (53.8%) and women (46.2%). In terms of age, the majority of the respondents (30.8%) were between 25 and 29 years old, 25.3% involved people between 18 and 24 and 24.2% were between 30 and 39. Finally, the vast majority of the respondents (86.9%) included graduates or post-graduate students.

4.2 Measures

The questionnaire was divided into two parts. The first part included questions on the demographics of the sample (age, marital status, gender, education). The second part included measures of the various constructs identified in the literature review section. Table 2 lists the operational definitions of the constructs in this theoretical model, as well as the studies from which the measures were adopted. The appendix lists the questionnaire items used to measure each construct. We employed a 7-point Likert scale anchored from 1 ("completely disagree") to 7 ("completely agree").

Table 2. Construct definition and instrument development

Construct	Operational Definition	Source
Privacy (PR)	Measuring the customers' privacy issues when using personalized services.	[26]
Trust (TR)	Measuring the customers' trust issues when using personalized services.	[27]
Happiness (HAP)	Measuring the customers' happiness when using personalized services.	[15]
Anxiety (ANX)	Measuring the customers' anxiety when using personalized services.	[15]
Intention to Purchase (INT)	Customers' intention to shop online based on personalized services	[28-29]

4.3 Data Analysis

Structural equation modeling was conducted using AMOS version 18.0 software, based on Byrne [32]. At first, a measurement model was created based on a confirmatory factor analysis, and then the structural model was built in order to test the hypothesized relationships.

Goodness of fit describes how well the model fits its data. Here, several fit indices were used to assess model-data fit. Root mean square error of approximation (RMSEA), comparative fit index (CFI) and $\chi 2/df$ ratio were all used to evaluate model-data fit (Byrne, 2009). RMSEA less than 0.05 suggests good model-data fit; between 0.05 and 0.08 it suggests reasonable model-data fit and between 0.08 and 0.01 suggests acceptable model data fit. CFI indices greater than 0.90 suggest good model-data fit and greater than 0.80 suggest adequate model-data fit. A $\chi 2/df$ ratio less than 3 is acceptable.

5 Findings

First, an analysis of reliability and validity was carried out. Reliability testing, based on the Cronbach alpha indicator, shows acceptable indices of internal consistency since all constructs exceed the cut-off threshold of 0.70. The AVE for all constructs ranges between 0.681-0.805, exceeding the cut-off threshold of 0.50. Finally, all correlations are lower than 0.80 and square root AVEs for all constructs are larger from their correlations. Our findings are illustrated in Table 3.

Table 3. Descriptive statistics and correlations of latent variables

					Construct				
Construct	**Mean**	**SD**	**CR**	**AVE**	**PR**	**TR**	**HAP**	**ANX**	**INT**
PR	5.35	1.63	0.923	0.813	**0.902**				
TR	3.05	1.44	0.896	0.745	-0.153*	**0.863**			
HAP	3.84	1.46	0.769	0.564	-0.120*	0.386**	**0.751**		
ANX	2.97	1.44	0.838	0.567	0.205**	-0.123*	0.167*	**0.752**	
INT	4.05	1.66	0.939	0.630	-0.203**	0.388**	0.601**	-0.161*	**0.794**

Note: Diagonal elements (in bold) are the square root of the average variance extracted (AVE). Off-diagonal elements are the correlations among constructs (all correlations are significant, **p< 0.01; *p<0.05). For discriminant validity, diagonal elements should be larger than off-diagonal elements. PR, Privacy; TR, Trust; HAP, Happiness; ANX, Anxiety; INT, Intention to Purchase.

The fit indices of the research model are presented on table 4. All values are within the recommended range. Specifically, $\chi 2/df = 1.96$, CFI = 0.96 and RMSEA = 0.07.

Table 4. Overall model fit indices for the structural model

Model fit indices	Results	Recommended value
χ^2/df	1.96 (x2 =; df =)	<= 3
CFI	0.96	>= 0.9
RMSEA	0.07	<= 0.08

The estimated path coefficients of the structural model were examined in order to evaluate our hypotheses. Figure 2 presents the analysis of the research model. Specifically, privacy has a positive effect on anxiety, supporting H2. On the other hand, privacy has no significant effect on happiness, rejecting H1. Regarding trust, a positive effect on happiness was found, supporting H3, while there was no significant effect on anxiety. Next, both happiness and anxiety were found to affect significantly intention to purchase. Happiness's effect on intention is positive, supporting H5, while anxiety's effect is negative, supporting H6. Square multiple correlations (R2)

are presented on figure 2 as well. The R2 for happiness was 0.43, for anxiety was 0.1 and that of intention to purchase was 0.85. Values higher than 0.26, imply high effect of the predictors of positive emotions and intention to purchase respectively.

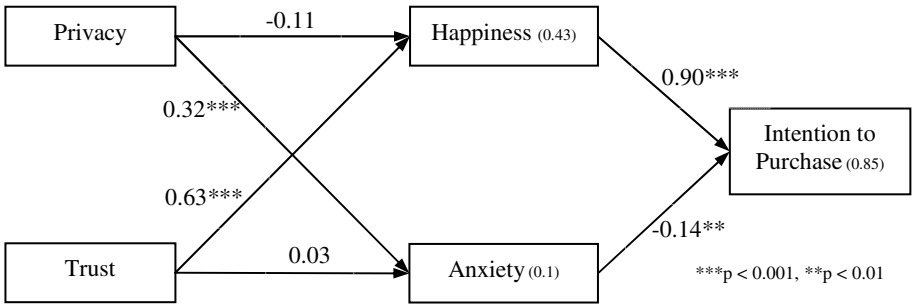


Fig. 2. SEM Analysis of the research model

6 Discussion and Conclusions

This study emphasizes on how the use of personalized services affects customers' intentions to proceed to online purchase. We investigate the effects of privacy and trust on the formulation of two critical emotional factors, namely happiness and anxiety, and the effect of the two emotional factors on purchase intention. Our results indicate that privacy and trust affect partially the emotional factors examined in this study. Specifically, privacy affects positively anxiety but has no effect on happiness. On the other hand, trust affects positively happiness but has no effect on anxiety. Moreover, happiness was found to have a very high positive effect on intention to purchase, while anxiety had a negative effect on intention to purchase.

Our study revealed that privacy issues have no effect on happiness, while a negative effect was expected, which may be due to the fact that the respondents shared their personal and private data willingly and might expect some privacy violations. Privacy was found to have a positive effect on anxiety, which was expected as for a customer with high privacy concerns, being unsure of how the information he shared with an online vendor are used, while it is the only way to use personalized services, makes him feel insecure, helpless and nervous. On one hand, our findings regarding the effects on happiness, which a positive emotion, contradict with previous studies that have found privacy to affect negatively one other positive emotion, that of enjoyment [20]. On the other hand, our findings regarding the effects on anxiety are consistent with previous studies that found high privacy concerns to be positively related with anxious customers [3]. In the case of trust, we found that it has a positive effect on happiness, but has no effect on anxiety. As expected, when users trust the online vendor are less worried about sharing their personal and private information and using the personalized services, which makes them feel more satisfied and excited. However, the insignificant effect of trust on anxiety was unexpected, which may be attributed to the fact that trust is not related with negative feelings, but it is the

factor of distrust that creates them [33]. Hwang and Kim [17] found a positive relation among trust and enjoyment, similar to our findings, but they found a negative relation between trust and anxiety, which contradicts with our results.

The basic emotions examined on this study, were both found to affect customers' intention to purchase. Happiness had a positive effect on intention as expected, because the more satisfied and excited are customers with personalized services, the more likely is that they will want to make purchases. Similarly, was found that anxiety has a negative effect on intention, as the more insecure, helpless and nervous a customer feels the less likely is that he will proceed to purchases. Our findings are consistent with Lu and Su [15], who found enjoyment to have a positive effect on intention and anxiety to have a negative effect on intention. Moreover, our results regarding personalized services as an Information Technology are consistent with Beaudry & Pinsonneault [6], who studying IT use found happiness to affect IT positively and anxiety to affect IT negatively,

Our empirical research has addressed several shortcomings of previous studies in the area. Schwaig et al. [3], studying customers' privacy, although they include anxiety in their research, they do not consider factors of positive emotions like happiness. Moreover, Lee and Cranage [23], examine personalization and privacy but do not consider any emotional factors in their study, factors which have been proved to affect both highly. Pappas et al. [20], while investigating the effects of privacy and enjoyment when using personalized services, they do not investigate any negative emotions. Positive and negative emotions should be studied together as they may have different effects on customers' behavior [34].

The present study is one of the few so far that explores customers' basic emotions while using personalized services for online shopping. We investigate how these emotions influenced by privacy and trust combine to affect customers' intention to purchase. Since this area is understudied, the proposed model shows an acceptable level of explaining how privacy, trust, happiness, anxiety and intention to purchase relate in order to fill this gap. Previous studies on online shopping have identified the importance of negative emotions over the positive. For example, a customer will be more influenced by reading a negative review for an online vendor than reading a positive one. However, using personalized services seems to change that, since positive emotions, and specifically happiness, are those that affect mostly customers' intention.

The findings of this study reveal that emotions affect customers' intentions to purchase through personalized services, with happiness having a much higher impact on intention. Hence, service providers should focus on customers' emotional responses in order to understand their intentions and change the way they communicate with their customers accordingly. Moreover, online vendors that offer personalized services should primarily target happiness over anxiety. Consequently, online vendors should try to increase customers' trust towards them since only trust was found to affect happiness.

As with any empirical study, there are some limitations. Firstly, in this study the sample used is highly educated (university and post graduates) and is expected to be familiar with online services. Also, our results are based on self-reported data so customers actually behavior might be different. Future studies should include other

methods; observation, interviews and web log file analysis in order to provide a better understanding of customers' experience. Moreover, we include in our study only the two basic emotions that have been identified from the literature as the most important. Following studies should include anger and sadness for a more thorough study. Finally, since privacy affects only anxiety and trust affects only happiness, future research should examine both privacy and trust as one factor and study their interrelation.

References

1. Zhou, L., Dai, L., Zhang, D.: Online shopping acceptance model - A critical survey of consumer factors in online shopping. Journal of Electronic Commerce Research 8, 41–62 (2007)
2. Lee, E.J., Park, J.K.: Online service personalization for apparel shopping. Journal of Retailing and Consumer Services 16, 83–91 (2009)
3. Schwaig, K.S., Segars, A.H., Grover, V., Fiedler, K.D.: A Model of Consumers' Perceptions of the Invasion of Information Privacy. Information & Management (2012), doi:10.1016/j.im.2012.11.002
4. Lee, D.J., Ahn, J.H., Bang, Y.: Managing consumer privacy concerns in personalization: a strategic analysis of privacy protection. MIS Quarterly 35, 423–444 (2011)
5. Penz, E., Hogg, M.K.: The role of mixed emotions in consumer behaviour: Investigating ambivalence in consumers' experiences of approach-avoidance conflicts in online and offline settings. European Journal of Marketing 45, 104–132 (2011)
6. Beaudry, A., Pinsonneault, A.: The other side of acceptance: studying the direct and indirect effects of emotions on information technology use. MIS Quarterly 34, 689–710 (2010)
7. Kuo, Y.F., Wu, C.M.: Satisfaction and post-purchase intentions with service recovery of online shopping websites: Perspectives on perceived justice and emotions. International Journal of Information Management 32, 127–138 (2012)
8. Kwon, K., Cho, J., Park, Y.: How to best characterize the personalization construct for e-services. Expert Systems With Applications 37, 2232–2240 (2010)
9. Adolphs, C., Winkelmann, A.: Personalization Research in E-Commerce - A State of the Art Review (2000-2008). Journal of Electronic Commerce Research 11, 326–341 (2010)
10. Adomavicius, G., Tuzhilin, A.: Personalization technologies: a process-oriented perspective. Communications of the ACM 48, 83–90 (2005)
11. Roberts, R.C.: Emotions: An essay in aid of moral psychology. Cambridge University Press (2003)
12. Zhang, H., Lu, Y., Shi, X., Tang, Z., Zhao, Z.: Mood and social presence on consumer purchase behaviour in C2C E-commerce in Chinese culture. Electronic Markets, 1–12 (2012)
13. Chiu, C.M., Wang, E.T.G., Fang, Y.H., Huang, H.Y.: Understanding customers' repeat purchase intentions in B2C e-commerce: the roles of utilitarian value, hedonic value and perceived risk. Information Systems Journal (2012)
14. Ethier, J., Hadaya, P., Talbot, J., Cadieux, J.: Interface design and emotions experienced on B2C Web sites: Empirical testing of a research model. Computers in Human Behavior 24, 2771–2791 (2008)
15. Lu, H.P., Su, P.Y.J.: Factors affecting purchase intention on mobile shopping web sites. Internet Research 19, 442–458 (2009)
16. Kay, R.H., Loverock, S.: Assessing emotions related to learning new software: The computer emotion scale. Computers in Human Behavior 24, 1605–1623 (2008)

17. Hwang, Y., Kim, D.J.: Customer self-service systems: The effects of perceived Web quality with service contents on enjoyment, anxiety, and e-trust. Decision Support Systems 43, 746–760 (2007)
18. Palvia, P.: The role of trust in e-commerce relational exchange: A unified model. Information & Management 46, 213–220 (2009)
19. Brandimarte, L., Acquisti, A., Loewenstein, G.: Misplaced confidences: Privacy and the control paradox. Social Psychological and Personality Science (2013)
20. Pappas, I.O., Giannakos, M.N., Chrissikopoulos, V.: Personalized services in online shopping: Enjoyment and privacy. In: 2012 International Conference on Information Society (i-Society), pp. 168–173. IEEE (2012)
21. Xu, H., Luo, X.R., Carroll, J.M., Rosson, M.B.: The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. Decision Support Systems 51, 42–52 (2011)
22. Dabholkar, P.A., Sheng, X.: Consumer participation in using online recommendation agents: effects on satisfaction, trust, and purchase intentions. The Service Industries Journal 32, 1433–1449 (2011)
23. Lee, C.H., Cranage, D.A.: Personalisation-privacy paradox: The effects of personalisation and privacy assurance on customer responses to travel Web sites. Tourism Management 32, 987–994 (2011)
24. Young, L.: Trust: looking forward and back. Journal of Business & Industrial Marketing 21, 439–445 (2006)
25. Martín, S.S., Camarero, C., José, R.S.: Does involvement matter in online shopping satisfaction and trust? Psychology and Marketing 28, 145–167 (2011)
26. Shen, J.: Social Comparison, Social Presence, and Enjoyment in the Acceptance of Social Shopping Websites. Journal of Electronic Commerce Research 13, 198–212 (2012)
27. Koo, D.M., Ju, S.H.: The interactional effects of atmospherics and perceptual curiosity on emotions and online shopping intention. Computers in Human Behavior 26, 377–388 (2010)
28. Ho, S.Y., Ho, K.K.W.: The Effects of web personalization on influencing users' switching decisions to a new website. In: PACIS 2008 Proceedings. Paper 67 (2008)
29. Ray, S., Ow, T., Kim, S.S.: Security Assurance: How Online Service Providers Can Influence Security Control Perceptions and Gain Trust. Decision Sciences 42, 391–412 (2011)
30. Chen, Y.T., Chou, T.Y.: Exploring the continuance intentions of consumers for B2C online shopping: Perspectives of fairness and trust. Online Information Review 36, 104–125 (2012)
31. Lu, Y., Cao, Y., Wang, B., Yang, S.: A study on factors that affect users' behavioral intention to transfer usage from the offline to the online channel. Computers in Human Behavior 27, 355–364 (2011)
32. Byrne, B.: Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming. Taylor & Francis (2009)
33. McKnight, D.H., Choudhury, V.: Distrust and trust in B2C e-commerce: Do they differ? In: 8th International Conference on Electronic Commerce: The New E-Commerce: Innovations for Conquering Current Barriers, Obstacles and Limitations to Conducting Successful Business on the Internet Conference, pp. 482–491. ACM (2006)
34. Barclay, L.J., Kiefer, T.: Approach or Avoid? Exploring Overall Justice and the Differential Effects of Positive and Negative Emotions. Journal of Management (2013)

Appendix

Measures		Loading
Privacy (PR)	**Personalization causes privacy problems:**	
PR1	Because it may keep track of my web behavior.	0.97
PR2	Because it may monitor my clicks and browsing records.	0.96
PR3	By exposing my personal information to unknown parties.	0.76
Trust (TR)	**The online vendor that offers personalized services:**	
TR1	Can be trusted at all times.	0.84
TR2	Can be counted on to do what is right.	0.89
TR3	Has high integrity.	0.86
Emotions	**In general, when I receive personalized services, I feel:**	
Happiness (HAP)	Satisfied	0.87
	Excited	0.84
	Curious	0.57
Anxiety (ANX)	Anxious	0.78
	Insecure	0.74
	Helpless	0.71
	Nervous	0.78
Intention to Purchase	**Based on personalized services:**	
INT1	In the future I intend to continue shopping.	0.93
INT2	My general intention to buy online is very high.	0.88
INT3	I will shop online in the future.	0.78

Trust and Privacy in the Shift from E-Commerce to M-Commerce: A Comparative Approach

Panagiota Papadopoulou[1] and Jean-Eric Pelet[2]

[1] University of Athens, Athens, Greece
peggy@di.uoa.gr
[2] IDRAC International School of Management,
LEMNA, IAE-IEMN – Nantes University, France
je.pelet@gmail.com

Abstract. Trust and privacy have been widely studied as key issues and success factors for e-commerce. The advent of m-commerce calls for revisiting these concepts and re-examining their antecedents in the mobile context. This paper attempts a comparative approach to the issues of trust and privacy in e-commerce and m-commerce. It investigates how trust and privacy are differentiated with the shift from the context of e-commerce to the context of m-commerce. Our analysis is supported by the results of an exploratory qualitative study in m-commerce.

1 Introduction

Even if year-over-year sales haven't increased dramatically since the fourth quarter of 2007 (ComScore, 2012), online commerce keeps on growing. Retail as an industry is experiencing important shifts in consumer behavior and the popularity of devices such as smartphones and tablets is having a significant impact on both online and in-store shopping. Some authors such as Serres (2012) or De Rosnay (2012) even talk of "symbiotical men" or "augmented men" or "little thumb" ("petite poucette"), when referring to the prosthesis such devices represent in our everyday life.

Trust and privacy, which have been recognized as issues of major importance in e-commerce, have been demised and retained in m-commerce. Activities such as purchasing online from a mobile site or leaving a comment on a social network, can involve more or less important risks linked to privacy and require trust. However, with the joint use of social networks with location-based applications, it seems that trust as well as privacy perceptions are differentiated. More and more people acquire smartphones, m-commerce is growing very rapidly, however, it remains difficult to explain why people may give away huge amounts of personal data from their mobile device, without being as reluctant or afraid of any privacy problems, whereas it can sometimes be more difficult to purchase from a "normal" e-commerce website.

The aim of this paper is to offer a comparative analysis of the concepts of trust and privacy in e-commerce and m-commerce. Trust and privacy have been extensively examined in e-commerce and remain prominent in m-commerce. However, the

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 50–60, 2013.

mobile environment and the mobile devices, in terms of their specific characteristics and the nature of the interaction they offer, require a re-examination of these two concepts and the factors that contribute to their establishment or to their lack of.

This paper will try to illustrate the differences which exist between the two contexts, e-commerce and m-commerce. A literature review comparing trust and privacy in e-commerce and m-commerce is presented in the next section. The paper continues with an analysis of the shift from e-commerce to m-commerce, around the pillars of trust and privacy. The section focuses on two directions, the added value offered in the mobile commerce setting compared to the conventional e-commerce one and the privacy paradox that emerges in the mobile context. In support of our analysis, the findings of an exploratory qualitative study related to trust and privacy in m-commerce are presented next. The paper ends with concluding remarks.

2 Literature Review

2.1 Trust and Privacy in E-Commerce

The lack of consumer trust in online shopping has been recognized as one of the major impediments for the success of e-commerce. Trust and its establishment have been related to consumer perceptions of privacy and security. Perceived privacy and security, in terms of handling consumer's private data, associated with a website has been found to influence trust (Flavian and Guinaliu, 2006). In a study of the adoption of social networking services, perceived privacy and perceived security were examined as the chief determinants of user trust (Shin, 2010). However, as the author notes, in the social networking services context, the effect of perceived privacy on trust is not as significant as showed by studies in the traditional e-commerce. The findings also include a strong effect of privacy on security.

Privacy is also the topic of a recent study (Schwaig *et al.*, 2012), which showed that consumers concern for information privacy is influenced by individual differences, namely self-esteem, consumer alienation and computer anxiety.

Privacy as a concept is closely intertwined with control. In e-commerce, there is a reported consumer demand for control over their personal data, in terms of disclosure or use. Disclosure of data would not aim at establishing intimacy but could be elicited by the potential of financial rewards, increased information and improved services as well as by the reputation of the vendor. Consumer demand for control also involves the ownership of the personal information collected and its use by vendors. The provision of control of personal data to consumers, together with the vendor's reputation can both lead to reduced perceptions of risk and in turn to trust (Olivero and Lunt, 2004).

Privacy has also been linked to personalization. In e-commerce, personalization refers to tailoring and recommending products and services according to specific consumer characteristics, based on knowledge about their preferences and behavior (Adomavicius and Tuzhilin, 2005). Personalization and privacy are linked with a paradoxical relationship. According to Lee and Cranage (2011), personalization implies customer disclosure of personal information, since rich data are needed for

delivering personalized offerings. In this sense, personalization can create privacy concerns, as customers feel that their privacy is invaded for the provision of personalized services. At the same time, personalization and positive perceptions of it, such as the perceived usefulness of personalized services, have a positive effect on customer willingness to share personal information.

The personalization – privacy paradox holds also in the mobile setting. Xu *et al.* (2011) examine the privacy personalization paradox in the context of location-aware marketing. The study is further analyzed in the next section.

2.2 Trust and Privacy in M-Commerce

Lee (2005) studied trust in the m-commerce environment with an interactivity orientation. Five interactivity components, perceptions of user control, responsiveness, connectedness, ubiquitous connectivity and contextual offer were shown to be factors influencing trust.

Yeh and Li (2009) found that customer trust towards a vendor in m-commerce is affected by customization, as a dimension of the mobile vendor's website quality, customer satisfaction towards the vendor and brand image, as a dimension of the mobile vendor quality. Their model also included the interactivity of the mobile vendor's website quality as well as the responsiveness of the mobile vendor's quality, however their effect on trust was not found to be significant.

Zhou (2011a) has examined trust in the context of mobile banking. His empirically tested model shows that trust is affected by structural assurance, information quality, trust propensity and system quality. Trust in turn affects perceived usefulness and usage intention. In Zhou (2011b), the author examined the critical success factors of mobile website adoption. Trust was found to be influenced by system quality, service quality and information quality of mobile websites. Trust was also an antecedent of satisfaction together with perceived ease-of-use and perceived usefulness.

However, in Suki (2011), an opposite relationship between trust and satisfaction is presented. The findings of the study reveal that users' trust in the vendor or the service/application of m-commerce is affected by their satisfaction with the vendor or the service/application of m-commerce. Similarly, Hung *et al.*'s (2012) work on repurchase intention in m-commerce found that trust in mobile shopping is influenced by satisfaction with mobile shopping. The same study showed that confirmation of expectations with mobile shopping is also a determinant of trust.

Siau *et al.* (2003) suggested that trust in m-commerce encompasses privacy of customer information and security of mobile transactions. They proposed a framework in which trust antecedents, apart from the factors that have been identified in e-commerce, such as vendor and website characteristics, include technology related factors. These involve technology of wireless services, technology of mobile devices and technology of mobile websites. Technology related factors are proposed to be the main barriers of trust in the conduct of m-commerce.

Xu *et al.* (2011) examine privacy in location-aware marketing, through the privacy calculus model. Personalization is positively associated with perceived benefits and

perceived risks of information disclosure, which are both antecedents of perceived value of information disclosure. In a more recent study, Zhou (2012) focused on location-based services and their usage intention. His results showed that privacy concerns are negatively related to trust.

A summary of the literature on mobile commerce related to trust and privacy is presented in Table 1.

Table 1. Summary of literature on trust and privacy in m-commerce

Study	Topic	Dependent variable	Antecedents	Consequences
Zhou (2011a)	Initial trust in mobile banking	Usage intention	Usage intention antecedents: initial trust, perceived usefulness. Trust antecedents: structural assurance, information quality, trust propensity and system quality	Perceived usefulness and usage intention.
Yeh and Li (2009)	Customer trust towards a vendor in m-commerce	Trust	Customization, satisfaction and brand image	
Siau *et al.* (2003)	Trust in mobile commerce	Trust	Vendor characteristics, website characteristics, technology of wireless services, technology of mobile devices, security framework	
Zhou (2012)	Usage intention of location-based services	Usage intention	Usage intention antecedents: performance expectancy, social influence, facilitating conditions, trust, perceived risk. Trust antecedents: privacy concerns	Perceived risk, usage intention
Zhou (2011b)	Mobile web sites adoption	Satisfaction	Satisfaction antecedents: perceived ease-of-use, perceived usefulness, trust. Trust antecedents: system quality, service quality and information quality	Perceived usefulness, satisfaction

Table 1. (*continued*)

Study	Topic	Dependent variable	Antecedents	Consequences
Lee (2005)	Interactivity and trust in m-commerce	Behavioral intention to use m-commerce	Behavioral intention to use m-commerce antecedents: trust, attitude towards using m-commerce and contextual offer. Trust antecedents: user control, responsiveness, connectedness, ubiquitous connectivity and contextual offer	attitude towards using m-commerce, behavioral intention to use m-commerce
Suki (2011)	Satisfaction and trust in vendors involved in mobile commerce	Trust in the vendor	Satisfaction with the vendor	
Hung *et al.* (2012)	Determinants of mobile shopping continuance	Continued intention toward mobile shopping	Continued intention toward mobile shopping antecedents: satisfaction, trust in m-shopping. Trust in m-shopping antecedents: satisfaction, confirmation	continued intention toward mobile shopping
Kao (2009)	Transaction trust on m-commerce adoption	Intentions to adopt m-commerce	Transaction trust (business trust, experience-based trust, security)	Intentions to adopt m-commerce
Xu *et al.* (2011)	Personalization-privacy paradox in location-aware marketing	Purchase intention	Purchase intention antecedents: Willingness to have personal information used in location-aware marketing. Perceived benefits and perceived risks of information disclosure antecedents: personalization	perceived value of information disclosure

3 The Shift from E-Commerce to M-Commerce

In e-commerce, trust formation has long been recognized as a demanding, continuous and complex process, which involves numerous elements from both a technical and business perspective. M-commerce inherits the issues that online vendors had to address to in traditional e-commerce environments. Trust and trust-related questions such as privacy and security remain important for m-commerce adoption (Siau and Shen, 2003). Despite the similarities between e-commerce and m-commerce and the common approaches toward the difficulties they share, trust formation in m-commerce can be different from trust formation in the traditional e-commerce context. This can be largely attributed to several factors which can be largely grouped into two categories, the specificities of the mobile devices and the specificities of the wireless telecommunications.

The specificities of the mobile devices include differences in the interface and a limited screen size, which sometimes imply limited features and functionality compared to those available in desktop and laptop computers. These characteristics related to mobile devices can be mapped onto perceptions of the ease-of-use and usefulness of using mobile websites.

Wireless telecommunication may also affect perceptions of security, which were inherited from e-commerce but are even more salient in the m-commerce setting. In a recent study of Lu *et al.* (2012) in the context of consumer-to-consumer e-commerce trust toward a platform was revealed to be a much stronger determinant of user satisfaction than perceived platform functionality.

3.1 The Mobile Added Value

Forrester Research reports that approximately half of all tablet users are using their devices for shopping purposes (2012). A smartphone can be more convenient than a laptop or desktop to get a product or service. Firstly because we have it everywhere at any time, making of us ubiquitarian citizen consumers. Secondly, and most importantly, smartphones own already stored information that enhances the speed of shopping, from the entrance in the online shop, to the payment.

The use of mobile devices for conducting online commercial transactions offers a number of possibilities that were not present in conventional e-commerce. Customers that need to have an actual look and feel of products before buying them now have the possibility to visit a store, see and experience the product of interest and then use their mobile device to buy it online. This is called "showrooming", an activity which is becoming common in brick-and-mortar stores, as reports show that 53% of mobile consumers use their device to "showroom" products (ComScores, 2012). This is particularly useful when products have a cheaper price online than in store.

Another possibility is using mobile devices for comparison shopping while being in a brick-and-mortar store, to find the lowest price or the best deals online. The capability to use mobile phones to search and shop online while in brick-and-mortar stores creates one more advantage for Internet sellers with the right pricing and products. Customers now compete with retailers even when they are in their store.

Convenience is important too since people want to be able to shop anytime. Daily specials can be an important way of gaining consumers. Indeed, new online shopping models such as daily deals and flash sales are generating excitement and impulse purchases. They are made easier if they can be done anywhere at any time. More and more consumers are saving money by shopping only when sellers provide discounts or coupons (ComScore, 2012) and over 35 percent of shoppers regularly look for deals online. A "mobile benefit" would therefore exist, with the embedded capability we now have to compare shops, looking for products, and purchasing from our smartphones and tablets. Consumers seek out and disseminate information about a product using social media and are not reserved about providing their views and opinions.

Among the elements that contribute to the massive use of mobile devices, is the growth in social media usage in shopping activities, also called "social shopping". Social networking is one of the most important factors in mobile shopping. Today

consumers are not passive shoppers, but they are active contributors to the shopping activity. They will abandon any company that does not serve their needs or is perceived as not giving value to them. If companies do not allow its customers to share items, or any other information, they may miss some customers that are used to communicate and get information from social network contacts before making any shopping decision. The instantaneous views on consumer profiles, m-commerce and the development of mobile tools (e.g. smartphones, tablets) reinforce the appetite consumers have for this "communication swiss-knife" style.

3.2 Privacy Concerns of Social Networks and Location-Based Applications: A Paradox

People live with their mobile devices, and mobile phones, in particular, have become a personal tool in the sense that people use them anywhere, anytime, for a range of tasks, for example to check their email or even set their alarm. Although, in general, people may use different computers at different places (e.g. at work, at home, etc), they usually have one mobile phone, which belongs to them and they are not willing to lend it. As such, using a mobile gives the impression that privacy is protected as it is a personal tool. Thus, people will probably take more risks with their mobile than with their computer.

Users are sometimes not aware that when they sign up for a social network service, they are giving away information to third parties, affiliates, partners, agencies, companies, etc. The latter can learn information that people would rather keep private and use it to target customized advertisements to them. By knowing prospect or customer location, at a specific time with the help of GPS, users could suffer from bombardments of advertising specifically targeted to them, not only via email or SMS/calls on their mobiles but also on public visual advertising i.e. assuming that public billboards will target ads specifically to customers through push GPS and social network applications.

All users' data is available on mobiles, in the applications to which users subscribe. Thus, companies have access to an incredible amount of data. This raises a question of ethics, regarding the use of the collected data. The data lifecycle problem must be seriously taken into consideration for m-commerce. Several issues may arise in this direction. Which data exist? How are they collected? How are they protected? Where are they stored? How are they used? Who are they shared with? How are they destructed?

The relative novelty of mobile devices could possibly explain why people are not really aware of their threats yet. The screen is smaller and it looks more intuitive and visual, thanks to the lack of keyboards for some of them, which can drive more spontaneous behaviors sometimes. There is no anti-virus on mobiles giving users the impression that it is safer than a usual computer. There is less prevention from risks in comparison to computers, probably because less important information is stored into a mobile device. This can develop trust towards m-commerce and m-philosophy in general, since a link between a fun aspect and social aspect is established while doing business, particularly with the use of social networks.

By downloading an application, the user will use a tool that fits perfectly to the screen and may think that it is simplified in comparison to usual computers. The simplicity of mobile applications allows users to share information much faster. It makes users not think about the consequences of sharing the information as they normally would, but rather wanting to get it out as fast as possible – "on the spot". It seems that self-disclosure of personal data is easier in mobile settings as convenience outweighs the risk of privacy. At the same time, there is a lack of applications on computers such as those that have been created for mobile devices, which facilitate the process of sharing information. Time pressure can entice users to prefer mobiles for such applications which are ready and faster to use than on computers. Transactions costs, in terms of time and money, are lower when using mobiles rather than computers because there are fewer stages. Customers thus prefer to use their mobiles for certain tasks.

Users being less risk-averse towards m-commerce could also be due to other features of the mobile context. It is easier to connect everywhere with an internet-enabled mobile device. Finally, participating in social networks via mobile devices and using location-based application in particular, is also a way to stay connected all the time with contacts.

4 Exploratory Study

An exploratory qualitative study has been conducted seeking to understand consumer perceptions and behavior related to mobile devices and social networks. According to our knowledge, research combining social networking services, m-commerce and trust is scarce, thus an exploratory approach of the topic, with a survey-based method using short interviews seemed appropriate. The main objective of this study is to investigate if the use of social networking applications on mobile devices is common, and if so, if users were influenced in shopping on m-commerce websites after using these social applications. In this direction and in an attempt to elicit data from consumers, the study has been conducted asking questions related to the following topics: use of social network systems on mobiles, use of mobiles for shopping, emotions and feelings felt following the visit of a mobile commerce website, factors affecting behavior with mobiles and social networking services and perception an "ideal" m-commerce website. Questions were structured and open, allowing for low duration interviews. Every interview, the duration of which ranged from 20 to 25 minutes, was re-transcribed offering a verbatim of around 80 pages corresponding approximately to 6 hours of recording.

Sample selection was primarily based on qualitative criteria. To follow the criteria of data saturation (Mucchielli, 1991, p.114), we interviewed 21 students. Students are deemed suitable as a sample even if their use has often been questioned in terms of their appropriateness. They share many characteristics with the profile of mobile and Internet users' population, such as age. As shown by several studies, Internet users tend to be young adults, while the Internet usage penetration within the age groups of 18–29 years raises up to 95% (Zickuhr, 2010; Pew Research Center, 2010). Hence,

although our sample presents a bias towards younger subjects, it can arguably be acceptable as representative of Internet and mobile users. In addition, our study benefits from the use of students since they are considered as an important group of online consumers (Delafrooz et al, 2010) and are useful as a sample for empirical studies in m-commerce, in line with previous research (e.g. Kim et al, 2008). We adopted a neutral attitude when interacting with them so as not to influence their answers. Participants were questioned without being allowed to look at their mobile phone. This was to ensure that they answered only using their memory to access the information reinforcing their use of the combination of social networking services and m-commerce websites of their choice. The sample was selected using a mix of age, gender and socio-professional background, in order to obtain a homogeneous sample.

On the topic of trust, when respondents are questioned about what encourages them to trust a particular social network/application and not another, we observe that peer recommendations represent an important reason why trusting an m-commerce website. Thereby, when "it is popular" (6/21) especially into the network of the respondent, when the "social network/m-commerce" application "has been heard from friends", 7/ 21 respondents seem more interested in the m-commerce website. On top of that, privacy concerns arise. "If it respects our privacy when sharing info" (4/21) and "when many other people trust it and have it" (5/21), respondents are inclined to install a social network application or visit an m-commerce website. Like for e-commerce, "when there is a special safe paying system like paypal" (3/21) respondents accept to pursue this initiative to reach a new m-commerce website. Finally, the application itself "*must be well-known*" by 2/21 respondents in order to trust it.

5 Discussion

The rapid growth of mobile commerce is associated with an increasing level of consumer experience with mobile devices. As mobile commerce grows to be an alternative shopping channel to traditional e-commerce, the effects of factors influencing online shopping may change accordingly (Zhou *et al.*, 2007). Trust and privacy remain high in the agenda of m-commerce success factors, both with a somewhat contradictory contribution, as enablers and at the same time as inhibitors of m-commerce. This dual role is especially evident when m-commerce is bundled with social networks and location-based applications.

As Xu *et al.* (2011) suggest, the conceptual structure of the personalization privacy paradox is context dependent. This can be deemed as the outcome of the contextualized value of the content delivered to customers in mobile environments, especially with location-based services. Personalization in location-based services involves tailoring content to individual customers' interests, location, identity, activity and time (Junglas and Watson, 2006). With the advent and widespread adoption of social networking and location-based services and applications, the need for personalization transforms to a need for a feeling of participation. The joint use of social networks and location-based services through mobile devices allows for ubiquitous presence and participation. This possibility can boost their self-esteem and create a feeling of

belonging to a community, where they can be constantly in touch with others. Consumers in order to leverage this new value inherent in these applications available anywhere, anytime seem to be willing to release personal data, probably with even less reluctance than for just having personalized content. In this sense, the shift from e-commerce to m-commerce also seems to signal a shift from the personalization-privacy paradox to a participation-privacy paradox.

Identifying mobile users with these characteristics is important for online vendors with a mobile presence. M-commerce websites and applications can achieve collection of data in exchange for consumers' feelings of active participation and social presence. In addition, Ranganathan *et al.*'s (2006) study on mobile users and their switching behavior showed that mobile users should be distinguished and treated differently according to their usage pattern of mobile services. Mobile users who use their mobile services for social, non-work related purposes are more inclined to switch providers than users who use the services more for functional and work-related needs. Therefore this user category should also be taken into account from mobile providers to ensure customer loyalty.

References

1. Adomavicius, G., Tuzhilin, A.: Personalization Technologies: A Process-Oriented Perspective. Communications of the ACM 48(10), 83–90 (2005)
2. ComScores, comScore Reports $44.3 Billion in Q1 2012 U.S. Retail E-Commerce Spending, Up 17 Percent vs. Year Ago, retrieved from the Internet the 13th of December. Published the 9th May (2012), at this address: http://www.comscore.com/Insights/Press_Releases/2012/5/comScore_Reports_44.3_Billion_in_Q1_2012_U.S._Retail_E-Commerce_Spending
3. Flavian, C., Guinaliu, M.: Consumer Trust, Perceived Security and Privacy Policy: Three Basic Elements of Loyalty to a Web Site. Industrial Management & Data Systems 106(5), 601–620 (2006)
4. Hung, M.-C., Yang, S.-T., Hsieh, T.-C.: An Examination of the Determinants of Mobile Shopping Continuance. International Journal of Electronic Business Management 10(1), 29–37 (2012)
5. Junglas, I., Watson, R.T.: The U-Constructs: Four Information Drives. Communications of AIS 17, 569–592 (2006)
6. Kao, D.T.: The Impact of Transaction Trust on Consumers' Intentions to Adopt M-Commerce: A Cross-Cultural Investigation. Cyberpsychology and Behavior 12(2), 225–229 (2009)
7. Lee, C.H., Cranage, D.A.: Personilisation – Privacy Paradox: The Effects of Personalisation and Privacy Assurance on Customer Responses to Travel Web Sites. Tourism Management 32, 987–994 (2011)
8. Lee, T.: The Impact of Perceptions of Interactivity on Customer Trust and Transaction Intentions in Mobile Commerce. Journal of Electronic Commerce Research 6(3), 165–180 (2005)
9. Lu, J., Wang, L., Hayes, L.A.: How do Technology Readiness, Platform Functionality and Trust Influence C2C User Satisfaction? Journal of Electronic Commerce Research 13(1), 50–69 (2012)

10. Olivero, N., Lunt, P.: Privacy versus Willingness to Disclose in E-Commerce Exchanges: The Effect of Risk Awareness on the Relative Role of Trust and Control. Journal of Economic Psychology 25, 243–262 (2004)
11. Ranganathan, C., Seo, D., Babad, Y.: Switching Behavior of Mobile Users: Do Users' Relational Investments and Demographics Matter? European Journal of Information Systems 15, 269–276 (2006)
12. de Rosnay, J.: Surfer la vie, Comment survivre dans la société fluide. Les Liens Qui Liberent (2012)
13. Seabright, P.: The Company of Strangers: A Natural History of Economic Life. Princeton University Press (2004)
14. Schwaig, K., Segars, A., Grover, V., Fiedler, K.: A Model of Consumers' Perceptions of the Invasion of Information Privacy. Information & Management (2012), doi:10.1016/j.im.2012.11.002
15. Serres, M.: Petite Poucette. Le Pommier (2012)
16. Shin, D.-H.: The Effects of Trust, Security and Privacy in Social Networking: A Security-based Approach to Understand the Pattern of Adoption. Interacting with Computers 22, 428–438 (2010)
17. Siau, K., Sheng, H., Nah, F.: Development of a Framework for Trust in Mobile Commerce. In: Proceedings of the Second Annual Workshop on HCI Research in MIS, Seattle, WA, December 12-13 (2003)
18. Siau, K., Shen, Z.: Building Customer in Mobile Commerce. Communications of the ACM 46(4), 91–94 (2003)
19. Suki, N.M.: A Structural Model of Customer Satisfaction and Trust in Vendors Involved in Mobile Commerce. Int. Journal of Business Science and Applied Management 6(2), 17–30 (2011)
20. Xu, H., Luo, X., Caroll, J., Rosson, M.B.: The Personalization Privacy Paradox: An Exploratory Study of Decision Making Process for Location-Aware Marketing. Decision Support Systems 51, 42–52 (2011)
21. Yeh, Y.S., Li, Y.-M.: Building Trust in M-Commerce: Contributions from Quality and Satisfaction. Online Information Review 33(6), 1066–1086 (2009)
22. Zhou, L., Dai, L., Zhang, D.: Online Shopping Acceptance Model – A Critical Survey of Consumer Factors in Online Shopping. Journal of Electronic Commerce Research 8(1), 41–62 (2007)
23. Zhou, T.: An Empirical Examination of Initial Trust in Mobile Banking. Internet Research 21(5), 527–540 (2011a)
24. Zhou, T.: Examining the Critical Success Factors of Mobile Website Adoption. Online Information Review 35(4), 636–652 (2011b)
25. Zhou, T.: Examining Location-based Services Usage from the Perspectives of Unified Theory of Acceptance and Use of Technology and Privacy Risk. Journal of Electronic Commerce Research 13(2), 135–144 (2012)

A Cloud Provider Description Schema for Meeting Legal Requirements in Cloud Federation Scenarios

George Kousiouris[1], George Vafiadis[1], and Marcelo Corrales[2]

[1] National Technical University of Athens, Greece
gkousiou@mail.ntua.gr, gvaf@iccs.gr
[2] Leibniz University of Hanover, Germany
corrales@iri.uni-hannover.de

Abstract. The advent of Cloud computing has created numerous significant challenges with regard to manipulation of data and especially personal data in cases of Clouds and federated Clouds. Existing legislation currently creates constraints and boundaries in the free usage of external Cloud providers. The aim of this paper is to provide a schema definition and usage mechanism (CPDS) that includes various levels of legal information that is necessary for automating the process of Cloud provider selection and data outsourcing. Thus the aforementioned constraints may be checked in an automated and machine understandable fashion and fully harvest the potential that is created by advances in Cloud computing like dynamic federation. In this direction, legal gaps and necessary actions are identified so that the automation avoids manual and bureaucratic steps that are necessary at the moment.

Keywords: Cloud computing, legal issues, personal data, data management, Cloud federation.

1 Introduction

With the advent of Cloud computing, new challenges have arisen with regard to the handling of personal data in infrastructures that are not under the direct control of either the end users or the application service providers, when using externalized infrastructure services from the IaaS/PaaS layer. This lack of control creates threats with regard to data accountability and legal action, especially with regard to current legislation[1].

On the other hand, the significant advantages of Cloud computing, like high availability, seemingly infinite resources, federation capabilities for exploiting multiple IaaS providers and pay-per-use business models are very intriguing for the business aspects of applications dealing with personal data. However, the "iron curtain" that lays in front of these providers with regard to how they manage their infrastructures (and is dictated by their business needs and confidentiality) hinders the available information that may be exposed and could alleviate the legal fears when it comes to personal data management.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 61–72, 2013.

For this reason, an intermediate solution should be found, that can expose critical legal information of the IaaS provider capabilities in order to meet legal constraints, while on the other hand not force the latter to share sensitive information regarding their infrastructure. This information should then be used in order to automate the negotiation process between the SP (Service Provider-the entity that is responsible for finding a suitable cloud resource and deploying the application service) and the IaaS provider (IP), or between IaaS providers in the case of dynamic cloud federation (case of resource sharing between IaaS providers). The aim of this paper is to provide such an XML-based definition of needed information and its usage in a dynamic multi-cloud utilization scenario, in order to bridge the gap between legal compliance and dynamic business models. The main triggering factor for this work is the OPTIMIS project's deployment and usage scenarios [13], that include federated and multi-cloud operations that may encounter legal boundaries and constraints [2].

The paper proceeds with introducing related work in Chapter 2. In Chapter 3, the main OPTIMIS deployment scenario is portrayed along with the legal constraints that prohibit its full potential. The structure of the necessary XML schema and mechanism is presented in Chapter 4 while its usage in an automated environment is portrayed in Chapter 5. The identified legal gaps that are necessary to complete the automation of the process in the mentioned scenarios are described in Chapter 6 while Chapter 7 concludes the paper.

2 Related Work

Legal implications and fears from the usage of Cloud computing have been highlighted in numerous research efforts. For example, the issue of losing control over data and the risks or concerns that this involves is highlighted in [3], as an aspect that is or will slow down the proliferation of Cloud usage. Jurisdictional issues, accountability and compliance are also critical in [4], and the usage of XML-based descriptions for legal text and rules has been taken under consideration, mainly in the SLA creation but not IP selection. In this interesting work, also the role of an intermediate third party trusted instance is considered, for the validation of auditing processes for example. LegalXML is an interesting attempt towards machine understandable contracts and terms, from which the following more interesting initiateves may be identified. LegalDocML[22] is a modeling language aiming at overcoming differences in XML-defined legal terms between different standards. LegalRuleML[23] is an ongoing effort to standardize legal rules and reasoning, while LegalXML eContracts is a recommendation towards standardized contract structures from a legal point of view. The latter may be used in conjunction to our work for describing for example the Standard Contractual Clauses. This family of recommendations may be valuable from an expression point of view, however it does not focus on provider selection.

On the other hand, IT scientists have devoted a lot of time and effort in order to enable from a technical point of view flexible and dynamic management [6] and sharing of resources between different Cloud providers in order to meet demand peaks,

minimize cost [20] or achieve higher availability (through multi-site placement) [20]. The federated aspects of Cloud computing model are considered capable to create immense potential as they offer significant performance gains as regards to response time and cost saving under dynamic workload scenarios [7].

3 OPTIMIS Federation Scenarios and Legal Constraints

In the OPTIMIS ecosystem [13], the Service Provider (SP) is an entity that needs to find a suitable Infrastructure Provider (IP) in order to deploy the application developer's service. In this operation, it must also meet the constraints set by the latter in the service level agreement that relate to the legal issues [14]. Furthermore, during operation, the IP may use external resources from other IPs in federated cloud scenarios, for various reasons such as cost minimization, risk management and avoidance of Service Level Agreement (SLA) failures due to overprovisioning. However, in that case the initial IP is assigned the role of assuring the legal aspects of data management by the federated resources. One of the most critical outcomes [11] is the determination of data location. The one who determines where the data are stored (in general this role is described as the Data Controller) is responsible in the end for meeting the legal requirements for data management manipulation. If the Data Controller decides to move the data to another provider, it must ensure that the proper agreements are in place prior to the federation. These agreements in the legal plane are summarized in the following cases:

- Binding Corporate Rules[8]: these refer to intra-company procedures (so in our case they apply to the case of a company located in Europe that wants to federate to its affiliate data center in e.g. Asia) for reassuring proper technical measures when it comes to data manipulation.

-Standard Contractual Clauses: these refer to inter-company agreements, with regard to how data are treated during the manipulation by the target federated provider. While these clauses have been formalized by the EU [15], their final form (that dictates the agreement between the home European IP and the target foreign IP) is subject to the agreement and/or possible modification from the involved IP parties.

-Intellectual Property rights: these refer to the ownership of the produced data from the utilization of a service offered by the IP.

The aforementioned documents (mainly the BCR and SCC) in order to be valid must be certified by the Data Protection Authority of the EU country of the originating (or home) IP. At the moment there is no automated way for performing the necessary legal checks so that the EU IP can on the fly check the legal compliance of the federated (external) IP. These checks may include the location of the federated IP, its security protocols and certifications, its main legal documents compatibility, such as the BCR, SCC and IPR declarations, with the EU law. If the target provider is located in the EU/EEA or the compatible countries [9], then the federation may be performed with minimal interventions. However if the IP is located in countries outside this list, then the federation cannot proceed if these legal checks are not in place. Thus it is

limiting the applicability of the federated scenarios to providers from the "white list" or includes the need for manual bureaucratic procedures prior to the federation.

Other requirements include the security from data loss, which is mainly reflected in data replication or multi-site placement techniques, and data encryption both at rest and during transfer. However these aspects are generally met in current technical solutions anyhow, due to well-established and pre-existing non-functional requirements of IT systems.

The complete scenario of this process appears in **Fig. 1**. In the OPTIMIS case, with different roles in different scenarios, we can predict that this role will be spread around multiple levels (SPs, IPs, federated IPs). However the main responsible is regarded the IP, since it receives the legal constraints and must act according to them. The home IP located in the UK may select from a variety of providers located in the EU/EEA and compatible countries. However they cannot exploit offerings from third countries, before the legal documents are certified by the respective DPA of the originating EU country. This certification process depends on each DPA and in many cases needs manual submission of the legal texts and validation.

4 Cloud Provider Description Schema Structure

In order to mitigate this effect and be able to have increased flexibility with regard to federation decisions, a suitable declaration procedure must be in place in order to implement this legal framework. In our case, this was decided to be implemented through a suitable XML schema structure, namely the Cloud Provider Description Schema (CPDS), that an IP should complete with its own information and make public, in order to be used during the selection process. Then the SP (or the IP when it is acting as an SP in the federated scenario) at every interaction should request this information, in order to filter the ones that do not meet the requirements. The information that needs to be included in such a description is detailed in the following paragraphs.

As a first step, the IPs must implement a way of declaring the locations of their data centers in terms of country of establishment. This location information may be the most critical, but it is only a part of the information that may be exposed by an IP regarding either legal information or in general capabilities information that is not confidential. These data must be suitably formatted in a machine understandable way so that they can be automatically processed and taken under consideration during a provider selection/ranking. The variety of information may include legal information (location of data, terms of service, legal notices etc.), security information (supported protocols, security services etc.), ecological information (percentage of renewable energy used, energy management certifications), resource and data management support (in terms of SLA guarantees, affinity rules support, available setting of replication etc.).

Not all of the above information is necessary in the legal context, but it was decided to be also included in order to explore also other aspects of IP selection (e.g. eco-efficiency, which indirectly may be linked to legal obligations deriving from the

Kyoto Protocol [10]). The higher levels of this schema appear in **Fig. 2**. The top level of the structure is the IaaSProviderType. The main type consists of a number of elements of more elaborate types.. The main subtypes are the following

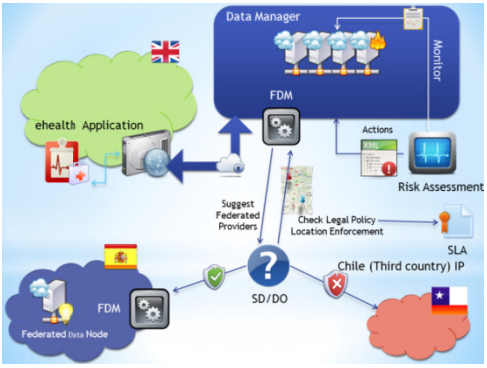


Fig. 1. OPTIMIS federation scenario: the home IP located in the UK may automatically select from a variety of providers located in the EU/EEA and compatible countries but not third countries even if the target IP is compliant but not adequately certified

- **LegalRequirementsType:** this element must be defined by an IP (**Fig. 3**), in order for the SP to be able to validate the legal adequacy of the former. In detail, the provider needs to insert the legal descriptions that govern its policies with regard to the SCCs, the BCR and the IPR with regard to added value from the services. For these textual descriptions it has been foreseen that they will be certified by an independent third party authority with expertise in the legal aspects of cloud computing (either public or private sector, similar to Verisign certification for website security). Another key characteristic refers to the ability of the provider to receive and fulfill requests regarding geographic location of service placement, which is another critical legal requirement, since the location of the data dictates the governing law. The security capabilities that are also critical from a legal point of view have been defined in the SecurityCapabilitiesType, given that they have also a strong technical aspect.
- **SecurityCapabilitiesType:** the provider must specify one element of this type, that contains information regarding its security features. These may include (**Fig. 3**) for example VPN support between the service VMs, Denial of Service attack detection capabilities, different authentication techniques (enumerated list with Boolean values that may be extended), encryption during data transfer or storage with a variety of strength options (bits used for the encryption) and/or a security certification by an external entity that has validated the security strategy of the company. This does not mean that the provider should publish the strategy itself, but just the ability, in order not to reveal sensitive information. Access control refers only to the IaaS level resources. In case of PaaS offerings, a similar field may be in the respective level.
- **DataManagementType:** the provider must specify one element of this type, that contains information regarding its data management features. These may include (**Fig. 5**) for example an element of DataReplicationType, stating the replication

strategy of the provider for data redundancy. These features may include a maximum replication factor, a configurable replication factor capability according to the client demands or a multi-site placement feature, which can lead in more reliable data storage. This is mainly dictated by the legal requirement for protection against data loss [19].

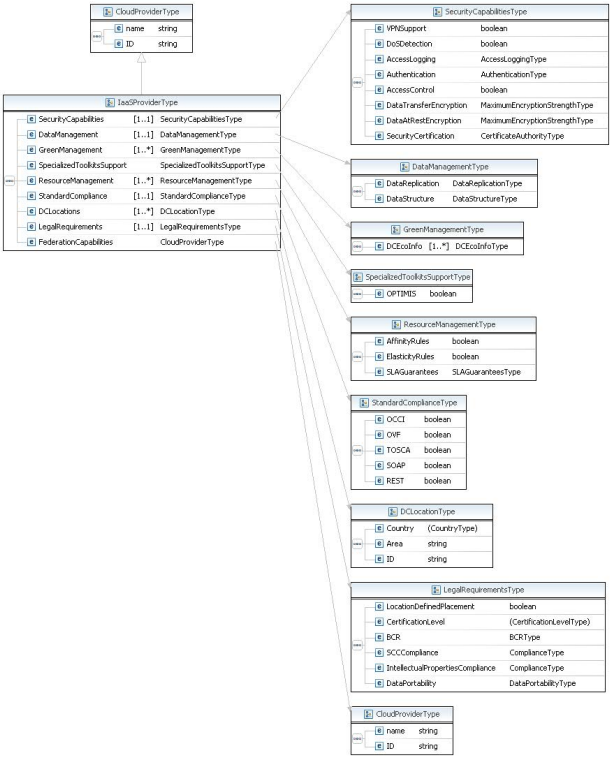


Fig. 2. Overall Structure of CPDS

- **GreenManagementType:** The provider must specify at least one element of this type, that contains information regarding its green management features per data center. The provider may declare different capabilities per data center (**Fig. 6**), given that not all of its infrastructures may be subject to green management, for example be equipped with energy generation from renewable sources. The basic subtype (**Fig. 7**) is the DCEcoInfoType, which contains information per DC, regarding the location (country, area and ID), the available green certificates (enumerated type consisting of the various available certificates) and the percentage of renewable energy produced (possibly on an annual average basis). The available certificates (as in all certificate types) may be signed by a CertificateAuthorityType, which is a generic type consisting of the authority's name and digital signature.

• **ResourceManagementType:** The provider must specify one element of this type, that contains information regarding its resource provision features per data center. These may include the possibility to have affinity rules (a feature that leads to optimized placement of interacting VMs), elasticity rules (a feature that leads to better self management capabilities) and the type of SLA guarantees (e.g. minimum availability)

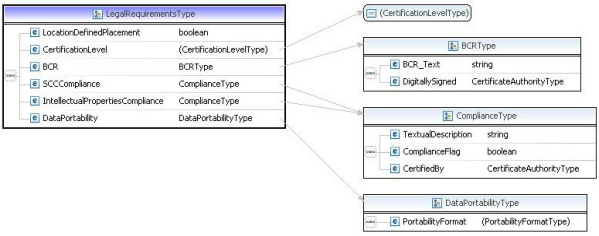


Fig. 3. Legal Requirements Type

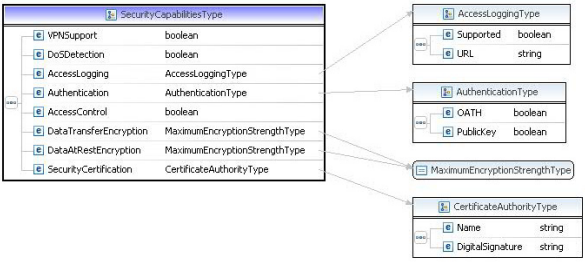


Fig. 4. Security Capabilities Type

• **DCLocations:** this element is used by the provider to declare at least one location of their data centers. This information is also used at the SP level, for the SP to filter out non-eligible locations (legal requirement). This type may include for example different locations for each DC, others in legal locations and others not. The final selection may take this under consideration, given that providers APIs usually offer a way of specifying desired target locations.

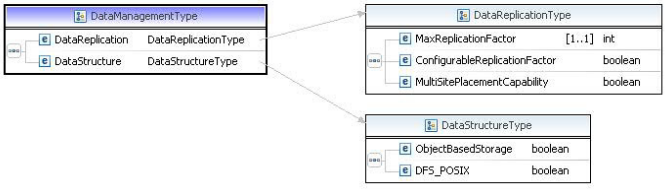


Fig. 5. Data Management Type

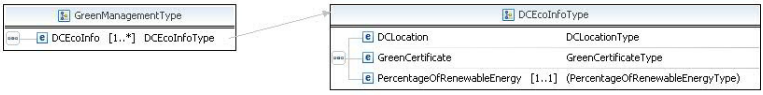


Fig. 6. Green Management Type

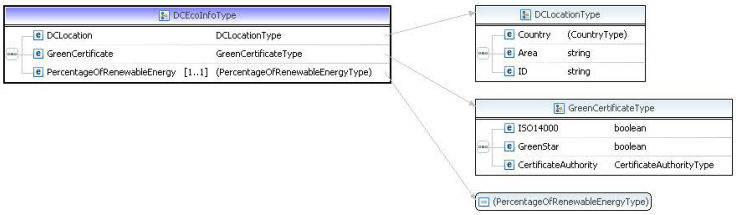


Fig. 7. Ecological Info Type

5 Usage during Runtime and Selection

The XML description of a provider following the aforementioned template is offered in the OPTIMIS platform through a restful GET interface of the Data Manager [12] component (getCPdescription). The description can be retrieved by any interested party (e.g. SP in initial deployment or the home IP in a federated scenario). Then the content of the XML may be directly checked through the checkLegal method in the OPTIMIS platform. This method compares the published information with the user requirements coming from the OPTIMIS SLA instance (e.g. location of provider, encryption capability and strength, eco-efficiency etc.) and concludes if the specific provider is eligible for the specific service.

In order to ensure that the level of information is valid, each field that requires an external verification from a certification authority needs to be digitally signed by the specific entity. Furthermore, the XML description itself should also be digitally signed, so that the requester of the document (in our case the SP or the home IP) may validate the content of the description. Furthermore, the caller should also validate that the digital signatures internally in the XML refer to the same entity as the provider that declares these capabilities and that the text itself is not altered and it is the same as the one submitted offline for validation in the DPA. This process appears in **Fig. 8** and is identical in concept to the validation process of websites through certification authorities such as Verisign etc.

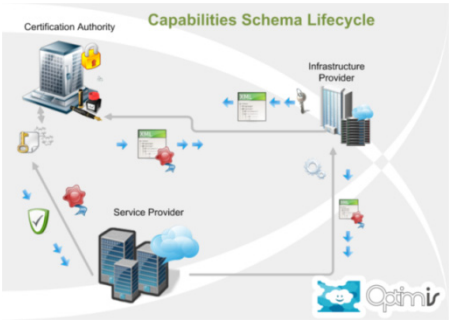


Fig. 8. Runtime Validation of An Infrastructure Provider's Description Schema

6 Identified Legal Gaps and Automation Steps

6.1 Validation of Arbitrary Terms and Conditions

Besides the improved selection of a provider based on this schema, the most important part is the legal section. In this, specific sections have been included like the BCR and SCC. These documents are arbitrary for each target company/provider and their conditions must be certified by the respective authority of each individual country (Data Protection Authority for the case of BCR and SCC) before a provider of this country federates to the target provider. This process is also manual in many countries, implying that before each federation decision for example, an employee of the home provider must go through a bureaucratic certification procedure of the target provider, bringing the text to the DPA and waiting for the resulting decision, that may be issued days afterwards. This of course creates an unrealistic legal framework for dynamic on the fly federation scenarios.

Automation Step

Thus a necessary automation step is that an EU country's DPA should act also as a certification authority for external providers that wish to be included in such federation operations. The process of certifying that an external (to EU) provider's BCR, and SCC framework is compatible to EU law should be handled offline between the this provider and the DPA and the latter should then act as a digital certification authority (similar to Verisign) and digitally sign the textual descriptions. This key then may be included in the specific provider's CPDI (Cloud Provider Description Instance). A provider interested in using the specific target provider for federation may then acquire the description and key, validate the key's correctness and thus acknowledge automatically that this provider is certified with relation to its policies.

6.2 Variability of IPR Strategies

The IPR strategies of each IP (or even at higher levels such as PaaS and SaaS) are determined by their strategy and potential exploitation/business schemes. However,

for the SP and the end user this is a very significant issue, if it affects their own exploitation or legal framework of operation, when choosing the Cloud provider. An example of this is the Google Docs Terms of Service, which specifies that: "When you upload or otherwise submit content to our Services, you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content. The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones."[17].. This may be against an exploitation scheme by the end user (and document producer) that includes confidentiality (such as a product design/documentation) and/or royalties for the material (e.g. book). Given that the terms and conditions for using services is in many cases arbitrary and based on company policy, their compliance validation in an automated framework means that they must be somehow grouped in classes and categories.

Automation Step

The categorization of IPR policies in predefined classes and the standardization of the latter create a suitable framework for automating the compatibility process between an end user and the provider. This would in essence be similar to the licensing schemes for software (GPL, BSD, Apache etc. licenses) and could also be accompanied with a rule-based decision framework that may examine cases of different classes compatibility (like in the case of [14] for licensing or even simpler approaches like an inter-license compatibility list) instead of strict "same class" comparison .

7 Conclusions

Dynamic business strategies (like Cloud federation, combination of private/public Clouds etc) that have emerged during the last years due to the technological breakthroughs in Cloud computing have created a number of issues with regard to the legal implications of data management for the application services.

In this paper, an effort to model these requirements in the form of a suitable description schema that the Cloud providers must complete and publish is portrayed, that may be used in an automated fashion during a legal check prior to the selection of a specific Cloud provider. The type of content regards information that normally is publicly available and/or can be found through relevant documentation. The content

Towards this direction, relevant legal and procedural gaps have been identified and corrective actions have been proposed (DPAs as certification authorities, IPR classes standardization need) in order to mitigate this problem and thus enable the usage of Cloud computing even in cases that involve personal data manipulation in dynamic and complex federation scenarios. DPAs acting as CAs may seem difficult, however this is a standard practice when it comes to internet certification and security.

For the future, we aim at enriching the XML schema with more details (e.g. by incorporating it in existing modeling approaches covering functional characteristics of

providers like PIM4CLOUD [18]) but also extending the legal checks implementation with the digital signature validation process.

Acknowledgement. This work has been supported by the OPTIMIS project and has been partly funded by the European Commission's IST activity of the 7th Framework Programme under contract number 257115.

References

[1] COM 11 final. Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), pp. 1–2 (2012), http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf

[2] OPTIMIS Project Deliverable D.7.2.1.3 "Cloud Legal Guidelines: Technical Implementation of Legal Requirements, Exploitation of the Toolkit in Use Cases and Component Licenses", pp. 23–33

[3] Chow, R., Golle, P., Jakobsson, M., Shi, E., Staddon, J., Masuoka, R., Molina, J.: Controlling data in the cloud: outsourcing computation without outsourcing control. In: Proceedings of the 2009 ACM Workshop on Cloud Computing Security (CCSW 2009), pp. 85–90. ACM, New York (2009)

[4] Pearson, S., Charlesworth, A.: Accountability as a Way Forward for Privacy Protection in the Cloud. In: Jaatun, M.G., Zhao, G., Rong, C. (eds.) CloudCom. LNCS, vol. 5931, pp. 131–144. Springer, Heidelberg (2009)

[5] Rochwerger, B., Breitgand, D., Levy, E., Galis, A., Nagin, K., Llorente, I.M., Montero, R., Wolfsthal, Y., Elmroth, E., Caceres, J., Ben-Yehuda, M., Emmerich, W., Galan, F.: The Reservoir model and architecture for open federated cloud computing. IBM Journal of Research and Development 53(4), 4:1-4:11 (2009)

[6] Buyya, R., Ranjan, R., Calheiros, R.N.: InterCloud: Utility-oriented federation of cloud computing environments for scaling of application services. In: Hsu, C.-H., Yang, L.T., Park, J.H., Yeo, S.-S. (eds.) ICA3PP 2010, Part I. LNCS, vol. 6081, pp. 13–31. Springer, Heidelberg (2010)

[7] Article 29 Working Party, WP 74, Transfer of Personal Data to Third Countries: Applying Article 26 (2) of the EU Data Protection Directive to Binding Corporate Rules for International Data Transfers

[8] European Commission, Commission Decisions on the Adequacy of the Protection of Personal Data in Third Countries, http://ec.europa.eu/justice/data-protection/document/international-transfers/adequacy/index_en.htm

[9] http://www.kyotoprotocol.com/

[10] OPTIMIS Project Deliverable 7.2.1.1 Cloud Legal Guidelines, pp. 91–93

[11] Kousiouris, G., Vafiadis, G., Varvarigou, T.: A Front-end Hadoop based Data Management Service for Efficient Federated Clouds. In: 2011 IEEE Third International Conference on Cloud Computing Technology and Science (CloudCom), November 29-December 1, pp. 511–516 (2011)

[12] Ferrer, A., et al.: OPTIMIS: A holistic approach to cloud service provisioning. Future Generation Computer Systems 28(1), 66–77 (2012) ISSN 0167-739X, 10.1016/j.future.2011.05.022

[13] Barnitzke, B., et al.: Legal Restraints and Security Requirements on Per-sonal Data and Their Technical Implementation in Clouds. In: Workshop for E-contracting for Clouds, eChallenges (2011)
[14] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:039:0005:0018:EN:PDF
[15] Gordon, T.F.: Analyzing open source license compatibility issues with Carneades. In: Proceedings of the 13th International Conference on Artificial Intelligence and Law (ICAIL 2011), pp. 51–55. ACM, New York (2011)
[16] http://www.google.com/intl/en/policies/terms/
[17] Brandtzg, E., Parastoo, M., Mosser, S.: Towards a Domain-Specic Language to Deploy Applications in the Clouds. In: Third International Conference on Cloud Computing, GRIDs, and Virtualization (CLOUD COMPUTING 2012), Nice, France, pp. 1–6 (July 2012)
[18] Djemame, et al.: Legal issues in clouds: towards a risk inventory. Phil. Trans. R. Soc. A 371(1983), 5, 20120075 (2013) (December 10, 2012) doi:10.1098/rsta.2012.0075
[19] Toosi, A.N., Calheiros, R.N., Thulasiram, R.K., Buyya, R.: Resource Provisioning Policies to Increase IaaS Provider's Profit in a Federated Cloud Environment. In: 2011 IEEE 13th International Conference on High Performance Computing and Communications (HPCC), September 2-4, pp. 279–287 (2011)
[20] Zhang, Z., Zhang, X.: A load balancing mechanism based on ant colony and complex network theory in open cloud computing federation. In: 2010 2nd International Conference on Industrial Mechatronics and Automation (ICIMA), May 30-31, vol. 2, pp. 240–243 (2010)
[21] OASIS LegalDocumentML TC, https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legaldocml
[22] OASIS LegalRuleML TC, https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legalruleml#technical
[23] OASIS LegalXMLeContracts TC, https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legalxml-econtracts#technical

Migrating into the Cloud: Identifying the Major Security and Privacy Concerns

Christos Kalloniatis[1], Vassilis Manousakis[1], Haralambos Mouratidis[2], and Stefanos Gritzalis[3]

[1] Department of Cultural Technology and Communication, University of the Aegean, University Hill, GR 81100 Mytilene, Greece
chkallon@aegean.gr, ct08081@ct.aegean.gr
[2] School of Architecture, Computing and Engineering, University of East London, Docklands Campus, 4-6 University Way, E16 2RD, London, U.K.
haris@uel.ac.uk
[3] Information and Communication Systems Security Laboratory, Department of Information and Communications Systems Engineering, University of the Aegean, GR 83200, Samos, Greece
sgritz@aegean.gr

Abstract. Cloud Computing is without a doubt one of the most significant innovations presented in the global technological map. The number of potential users enrolling and using cloud services increases exponentially on a daily bases. The great demand from online users for cloud services along with the reduced operational costs that the latter offers has motivated many organisations and companies to consider migrating organizational services, data and applications on the Cloud. However, despite the various positive characteristics of all cloud service models like reduced costs, better availability insurance, on demand data storage and computation power, cloud migration may hinder various security and privacy issues that need to be considered before an organization or company decides to move on to the Cloud. This paper aims to reveal the major security and privacy concepts for all three cloud service models and through a systematic analysis to guide the organization's stakeholders in deciding which service model best fits their needs based on their security and privacy requirements.

1 Introduction

According to National Institute of Standards and Technology (NIST), Cloud Computing delivers three different types of services to the end users that derive from three different models. The delivery models are IaaS, PaaS and SaaS, each one of them providing three distinct types of resources, like virtual infrastructure resources, application platforms and software services. Each delivery model is considered as separate layer that is depended from each other and with IaaS being the foundation, PaaS sits on top of IaaS and SaaS sits on top of PaaS. So, as the end users combine different type of services for implementing their requirements, they need to consider the various combined security and privacy threats that are behind these services.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 73–87, 2013.

Another factor that should be considered is the impact of deployment model on privacy and security measures. Security and privacy risks seem to have larger impact on public, hybrid and community cloud, compared to the rest of the deployment models. On the other hand, cloud consumers should keep in mind that while private cloud deployments are theoretically safer, still the same security and privacy threats do apply. The only difference is the users' group that the model is based on. In this deployment model, users starting from the administrator down to the simple one are trusted, but this does not imply a more secure and privacy oriented model.

In general, the more low level services the client requests the more responsible for security and privacy is, but still the cloud vendor has an important role on managing and implementing security and privacy measures even in low levels of abstraction.

The various innovations that cloud computing introduced in its operational environment vary from the traditional "trusted" environment where today's information systems rely on. These innovations hinder new security and privacy concepts that need to be identified in order to protect the traditional systems when migrating to cloud environments. This is exactly the scope of the specific paper. Specifically, in section 2 the critical cloud threats are presented as they have been examined from the respective literature. At the end of section two an analysis of the respective threats along with the cloud critical areas and cloud service models is conducted. In section 3 the description of the newly identified security and privacy concepts are described along with respective case diagrams. At the end of section 3 a matching between the newly concepts and their applicability on the cloud service models is presented in order to reveal the linkability and influence on every concept on the respective service model. This matching can also assist in informing the cloud users about what they should consider when migrating to the cloud as far as security and privacy requirements are considered, by matching them with each cloud service model separately without omitting of course to make a solid research for the above mentioned scenarios that meet their needs.

2 Critical Threats in Cloud Computing

In order to identify the major security and privacy cloud concepts an analysis of the major threats in cloud computing was conducted. The threats identified are based on the identified issues of the Cloud Security Alliance (CSA) report [1] as well as the ones from Gartner presented in [2] and mentioned in [3].

2.1 Threat #1: Abuse and Nefarious Use of Cloud Computing

Abuse and nefarious use of cloud benefits derive from the result of several reasons. For example, the constant advertisement of cloud's advantages result in attraction of more and more users in order to test their services, only to make cloud a giant pool of potential victims and attackers that want to exploit cloud vulnerabilities or even use cloud's computing power to perform illegal activities, all the above combined with Inadequate identity management[1] and lack of know-how[2] converts cloud from a

[1] Partial anonymity through weak registration.
[2] Limited of fraud detection capabilities.

ubiquitous and convenient resource pool, into an unsafe place to migrate someone's business vital operations. Several examples of this kind of usage are hosting of Zeus botnet, Trojan horses, Microsoft and Adobe PDF exploits, etc. The specific threat is matched with data center operations and Incident Response, Notification and Remediation domains and has applicability on IaaS and PaaS service models.

2.2 Threat #2: Insecure Interfaces and APIs

A variety of software interfaces and APIs are in use in order for the cloud services to be managed by the customers. Several actions like, management, provisioning, orchestration and monitoring are carried out through them. Customers, organizations and third parties interact with general cloud services through APIs, in order to process in implementing various services for their customers. Thus security and mainly availability of these APIs are of critical importance in the cloud environment. Confidentiality, integrity, availability and accountability are some of the issues that organizations are exposed through vulnerable APIs and interfaces[3]. The specific threat is matched with the application security domain and has applicability on IaaS, PaaS and SaaS service models.

2.3 Threat #3: Malicious Insiders

A malicious insider is a realistic scenario that a client cannot take immediate action. Opaque processes and procedures, not strict access to cloud's resources both physical and virtual, deficient monitoring, policy incompliance and improper employee hiring standards and in general lack of transparency are creating an attractive environment that could enable a potential adversary to gain control over cloud services and tamper data that rely on them. The specific threat is matched with the Governance and Enterprise Risk Management as well as with the Traditional Security, Business Continuity and Disaster Recovery domains and has an applicability on IaaS, PaaS and SaaS service models.

2.4 Threat #4: Shared Technology Issues

Virtualization is the concept that cloud computing notion was build. Dynamic provisioning of services in multi-tenant environment due to hardware virtualization (e.g., CPU,GPU, RAM, Disk partitions etc.) are promising advantages. On the other hand, the underlying infrastructure does not offer strong isolation between tenants, and as a result a virtualization hypervisor was implemented to fill this gap but still the issue has not completely addressed. As a matter of fact there are several examples like Joanna Rutkowska's Red and Blue pill and Kortchinksy's CloudBurst presentations that impact other tenants operations, tamper with data, network traffic, etc. The specific threat is matched with the data center operations and virtualization domains and has applicability on IaaS service model.

[3] Hidden filed manipulation, reusable tokens or passwords, clear-text authentication or transmission of content, improper authorizations, etc.

2.5 Threat #5: Data Loss or Leakage

Due to cloud's architecture the threat of data compromise increases. Data loss or leakage (through virtualization flaws) can cause unrecoverable damage and serious implications[4]. Insufficient AAA controls and encryption and software keys, system and operational failures, data lifecycle management challenges, compliance issues, vendor and client reliability are examples that derive from this threat. The specific threat is matched with the Information management and data security, the encryption and key management as well as the identity and access management domains and has applicability on IaaS, PaaS and SaaS service models.

2.6 Threat #6: Account or Service Hijacking

Phishing frauds, vulnerabilities exploitation, software exploitation or even user's personal choices (reused passwords) are methods that can lead to various threats in a cloud environment as for example hijacking. The damage that could cause a breach in terms of eavesdropping, tampering, service confidentiality, integrity and availability, is great. The specific threat is matched with the governance and enterprise risk management, the incident response, notification and remediation as well as the identity and access management domains and has applicability on IaaS, PaaS and SaaS service models.

2.7 Threat #7: Unknown Risk Profile

Seemingly insignificant factors about security should be considered by organizations. Software versions, updates, compliance, security practices and design, log files, information about the co-tenants, maintenance, who has access to the data or who is responsible or what data will be disclosure in case of an incident, how the data are stored in case of an incident, etc., all the above mentioned constitute an Unknown risk profile that companies should carefully weight. The specific threat is matched with the governance and enterprise risk management, the legal issues: contracts and electronic discovery, the data center operations as well as the incident response, notification and remediation and has applicability on IaaS, PaaS and SaaS service models.

2.8 Threat #8: Privileged User Access

Migrating to a cloud solution may result in loss of physical control over the organization operations and functions. Concerns as far as, "who" has access to data and the procedures in general, which are the hiring requirements, which is the level of access are posed. The specific threat is matched with the governance and enterprise risk management, compliance and audit and identity and access management domains and has applicability on IaaS, PaaS and SaaS service models.

[4] Brand and reputation damage to compliance violations and legal ramifications, etc.

2.9 Threat #9: Regulatory Compliance

Cloud providers are obliged to follow laws and regulations of each country where respective cloud services are reside from. Each country has different regulations as far as certain[5] procedures are done and the customer should be completely aware of them only to take his decisions. The specific threat is matched with the governance and enterprise risk management and compliance and audit domains and has applicability on IaaS, PaaS and SaaS service models.

2.10 Threat #10: Data Location

Security, privacy and data lifecycle procedures are strictly related to the country that the respective cloud services reside from. For example large datacenters may reside on foreign countries that have different jurisdictions specifications and regulations compared to the client's country. Client should be aware of that and make explicitly clear to the vendor the demands they have in mind. The specific threat is matched with the governance and enterprise risk management and compliance and audit and legal issues domains and has applicability on IaaS, PaaS and SaaS service models.

2.11 Threat #11: Lack of Data Segregation

Multi-tenancy in cloud computing is a basic concept that raises questions about the level of isolation between the tenants. Data should be completely isolated through the entire data lifecycle in order for the client to be protected. The specific threat is matched with the encryption and key management and virtualization domains and has applicability on PaaS and SaaS service models.

2.12 Threat #12: Lack of Recovery

In case of a disaster a solid recovery system should be in preparedness, just to restore services and data in their previous healthy state. The specific threat is matched with the traditional security, business continuity and disaster Recovery, incident response, notification and remediation domains and has applicability on PaaS and SaaS service models.

2.13 Threat #13: Investigate Support

In case of a security violation a properly configured forensics system should be ready, in order to examine the causes and the circumstances of the incident. Such actions are difficult due to cloud's nature, but provider should be ready to deal with this kind of emergencies. The specific threat is matched with Security as a Service and incident response, notification and remediation domain and has applicability on IaaS, PaaS and SaaS service models.

[5] Data processes, security and privacy procedures, etc.

2.14 Threat #14: Long-Term Viability

Cloud provider should have safety measures in case that something breaks tis service continuity (bankruptcy, DDoS attacks, etc.). Customer's data not only should be available in those situations, but there should be in their last healthy state. The specific threat is matched with portability and operability and traditional security, business continuity and disaster recovery domains and has applicability on IaaS, PaaS and SaaS service models.

CSA has also issued a report [4] regarding the cloud critical areas, with Cloud Architecture included, that Cloud providers and users should take into account, related both to strategic and tactical Security and Privacy pain points that exist in a Cloud environments and falls into any combination of cloud service and deployment model. The domains are divided into two main categories, governance and operations. Strategic and policy issues are addressed through governance domains, while the operational domains deal with security concerns and implementation techniques within the Cloud architecture. The critical areas are: a) Cloud Computing Architectural Framework, b) Governance and Enterprise Risk Management, c) Legal Issues : Contracts and Electronic Discovery, d) Compliance and Audit, e) Information Management and Data Security, f) Portability and Operability, g) Traditional Security, Business Continuity and Disaster Recovery, h) Data Center operations, i) Incident Response, Notification and Remediation, j) Application Security, k) Encryption and key management, l) Identity and Access Management, m) Virtualization and n) Security as a Service.

The matching between the aforementioned threats and the respective cloud service models is presented in table 1. In table 2 a matching between the aforementioned threats and the cloud critical areas is presented. These matching can assist the developers in identifying the possible drawbacks that will have to solve when deciding a cloud migration on a specific cloud model. Also this analysis assisted in the basic security and privacy concepts identification presented in section 3.

3 Security and Privacy-Oriented Concepts

This section describes the basic security and privacy properties that constitute the basic issues that need to be considered when a migration to the cloud needs to be established. The aim of this section is twofold. Firstly it aims on revealing and describing a number of concepts some of which are derived from the respective literature.

Secondly it aims on identifying the applicability of every concept on the respective cloud service model thus assisting the stakeholders on deciding which security and privacy properties need to be realised in order to satisfy their own goals on every cloud service model respectively. The concepts proposed are mainly derived from the European Commission Draft Report on Security Issues in Cloud Computing [5] as well as from our previous work presented in [6-13]. However, new concepts are also introduced and explained in order to form a complete set for covering all the respective cases.

Table 1. Matching Major Cloud Threats with Cloud Service Models

	IaaS	PaaS	SaaS
Threat #1: Abuse and Nefarious Use of Cloud Computing	x	x	
Threat #2: Insecure interfaces and APIs	x	x	x
Threat #3: Malicious Insiders	x	x	x
Threat #4: Shared technology issues	x		
Threat #5: Data Loss or Leakage	x	x	x
Threat #6: Account or Service Hijacking	x	x	x
Threat #7: Unknown Risk Profile	x	x	x
Threat #8: Privileged user access	x	x	x
Threat #9: Regulatory Compliance		x	x
Threat #10: Data Location	x	x	x
Threat #11: Lack of Data Segregation		x	x
Threat #12: Lack of Recovery		x	x
Threat #13: Investigate Support		x	x
Threat #14: Long-term Viability	x	x	x

Table 2. Matching Threats with Critical Cloud Areas

	Gov. & Enterpr. Risk Manag.	Legal Issues	Compliance and Audit	Inform.Manag & Data Security	Portability and Operability	Traditional Security, Business Continuity & Disaster Recovery	Data Center operations	Inc. Resp., Notific. & Remediation	Application Security	Encryption and key management	Identity and Access Management	Virtualization	Security as a Service
Threat #1: Abuse and Nefarious Use of Cloud Computing							x	x					
Threat #2: Insecure interfaces and APIs									x				
Threat #3: Malicious Insiders	x					x							
Threat #4: Shared technology issues							x					x	
Threat #5: Data Loss or Leakage				x						x	x		
Threat #6: Account or Service Hijacking	x							x			x		
Threat #7: Unknown Risk Profile	x	x					x	x					
Threat #8: Privileged user access	x					x							
Threat #9: Regulatory Compliance	x	x	x										
Threat #10: Data Location	x	x	x	x									
Threat #11: Lack of Data Segregation				x		x							
Threat #12: Lack of Recovery						x		x					
Threat #13: Investigate Support								x					x
Threat #14: Long-term Viability	x					x							

3.1 Isolation

The specific concept is referred to the complete seal of user's data inside the Cloud computing environment. Cloud computing resources sharing among a multi-tenant environment, poses the risk of any kind of information disclosure. As a result strong isolation must be achieved inside the cloud environment. Isolation is meant to address data disclosure in two ways, firstly, from purpose limitation point of view and secondly from the aspect of hypervisor hardening [5].

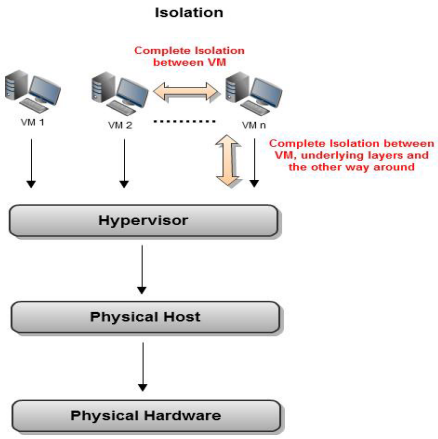


Fig. 1. Isolation Example

3.2 Provenanceability

The specific concept is referred to a vm's provenance mapping. Building a virtual machine's background tree makes it easier to get information about its parent image. The goal is to gather information about the reason of creating a new image, modifications, updates, vulnerabilities, etc. inside the cloud environment. The above, can be used to trace malicious actions of illegal content inside the vm image or let the owners know of a derived image that the parent image had for example a security problem [14].

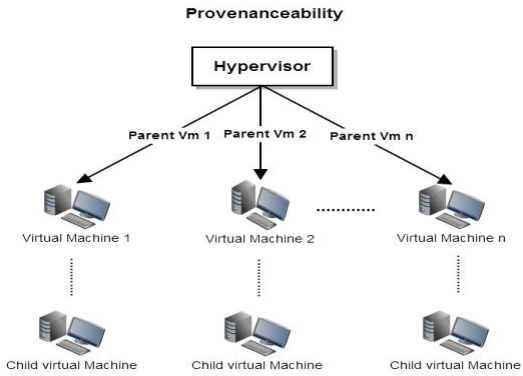


Fig. 2. Provenaceability Example

3.3 Traceability

Traceability concept aims to give the ability, for the data to be traced or not by the user. Data erasure is a major problem in web-based systems and still continues to

exist in clouds. Many cases have been documented for privacy violation due to improper data deletion (documents, photos, etc.). The traceability concept aims to reassure the clients that their data have been completely deleted or stay invisible and anonymized through the ability of tracing them among the data repositories.

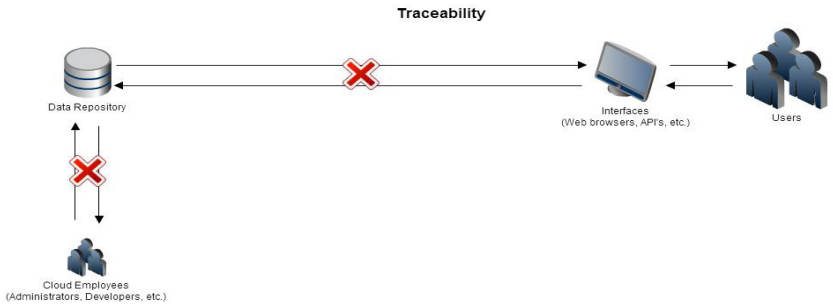


Fig. 3. Traceability Example

3.4 Availability

The specific concept tends to protect the system through specific procedures and techniques from meeting the above unwanted situations. Migration of enterprises operations to the cloud means that cloud provider is obliged to provide continuously the services to the specific enterprise no matter what. DoS attacks, physical disasters and hardware failure may result in breaking the continuity of service providing. That concept is to heal through specific procedures and techniques the above unwanted situations.

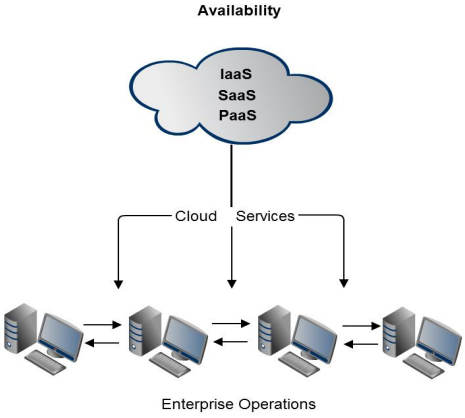


Fig. 4. Availability Example

3.5 Integrity

Integrity is referred to the fact that client's data remains in the exact form in which he send it in the first place. According to the EU directive [5], cloud providers must assure the user that his data have not got tampered while they were passing through the whole data life cycle. Integrity concept aims to provide to user assurances that their data have not been tampered somehow.

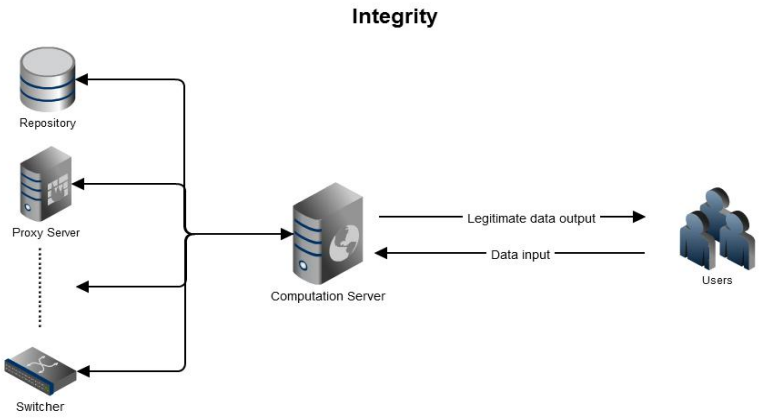


Fig. 5. Integrity Example

3.6 Confidentiality

Confidentiality issue derives from the fact that cloud is multi-tenancy environment and many of its resources are shared. That raises concerns about the data that travel inside the cloud, from the cloud provider to the client, and vice versa. Encryption techniques and authorization and authentication mechanisms, ensures that data's confidentiality is preserved [5].

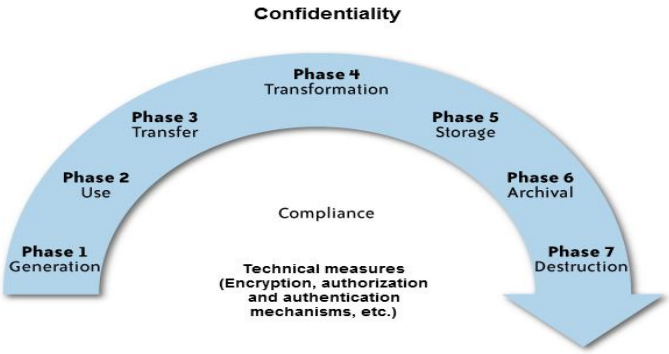


Fig. 6. Confidentiality Example

3.7 Transparency

Transparency concept is referred to the cloud vendor's obligation, to be completely clear about their procedures and functions. In order to preserve integrity and confidentiality of a client's data, transparency in several areas of cloud's procedures should exist. According to the EU directive, transparency must exist in regard to the contractors and subcontractors that cloud providers are related to and the internal cloud operations and procedures that the provider follows in certain circumstances [3, 5].

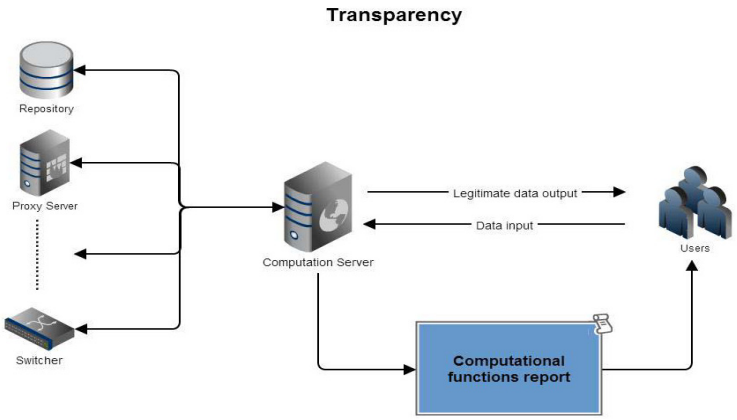


Fig. 7. Transparency Example

3.8 Intervenability

Intervenability concept is referred to the fact that, the users should be able to process their data despite the cloud's service architecture. A cloud vendor may rely on other provider's (subcontractor) services in order to offer his services. That should not be an obstacle for the user to intervene[6] to his data, in fact cloud vendor must be able to provide all the technical and organizational means to this goal including subcontractors [5].

3.9 Portability

Portability concept aims to achieve transferability as data are concerned, among different cloud providers and services. As we mentioned earlier data or vendor lock-in could result in lack of data portability and interoperability between different cloud services. The use of a standard format could impose obstacles in the transfer of personal data or even result in data disuse, due to the lack of compatibility, if a cloud vendor is bankrupted [5].

[6] Access, rectification, erasure, blocking and objection.

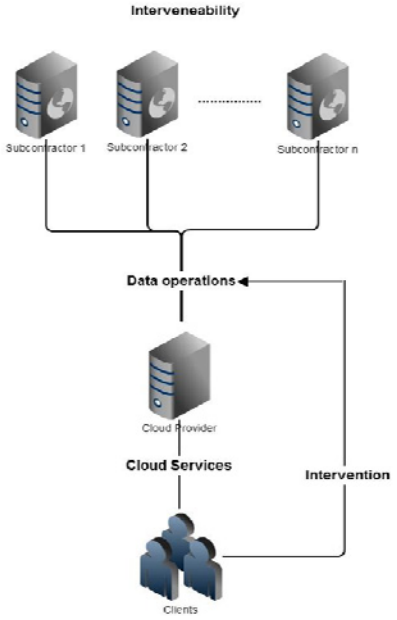


Fig. 8. Intervenability Example

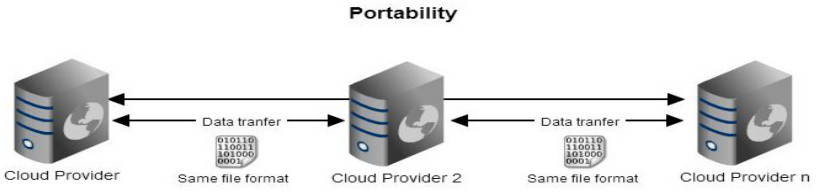


Fig. 9. Portability Example

3.10 Accountability

Accountability concept is referred to the fact that, cloud providers should provide information anytime about an incident. The cloud architecture[7] makes a complex form of an informational system. In terms of management and audit controls, this fact could result in very difficult manageability of incidents. A cloud provider should be able at any time to provide information about what an entity did and when, just to trace malicious actions from the whole cloud infrastructure [5].

[7] International services residual.

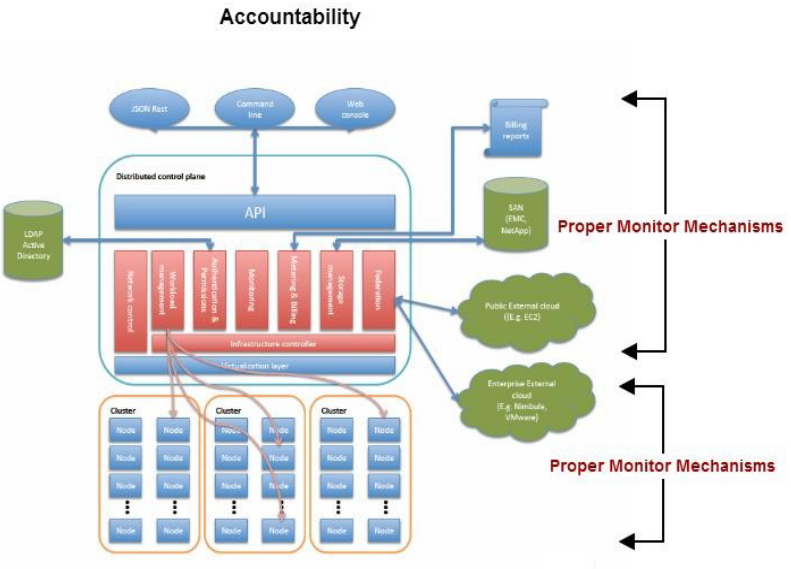


Fig. 10. Accountability Example

It should be mentioned that beside the above presented concepts other security and privacy properties do exist that can be taken under consideration when considering migrating an Information System to the Cloud. In [6-13] a number of security and privacy concepts are presented. These concepts have been presented for transforming privacy from a generic, mostly legal concept, to a technical requirement that analysts can enroll in the analysis and design process. These concepts also play a significant role in the migration process but due to space limitations are not presented analytically in this paper.

In the following table a matching between the aforementioned concepts and the cloud service models is presented. Based on the following table analysts can identify which are the security and privacy concepts belonging to their system and how these

Table 3. Matching Security and Privacy Properties with Cloud Services Models

	IaaS	SaaS	PaaS
Property #1: Isolation	x	x	x
Property #2: Provencability	x		
Property #3: Traceability		x	
Property #4: Availability	x	x	x
Property #5: Integrity		x	
Property #6: Confidentiality	x	x	x
Property #7: Transparency		x	
Property #8: Intervenability	x	x	x
Property #9: Portability		x	x
Property #10: Accountability	x	x	x

concepts can constitute an initial obstacle during the migration of a traditional system on a cloud environment. Usually when analysts consider the cloud migration their main goal is to decide on which service model they are interested in migrating to. The identified concepts and the respective matching is a start for creating a holistic process for assisting analysts on receiving the proper decisions.

4 Conclusions

Cloud computing environments offer great degree of scalability, flexibility and resource pooling thus elevating its use leading to its great expandability and applicability noted nowadays [15]. Many users, private companies and public organisations on a daily basis consider migrating their systems into the cloud in order to take advantage of these possibilities. However, cloud innovations hinder new undiscovered threats that analysis and users need to be considered before deciding the big change of cloud migration. This paper is an initial step on identifying a set of new concepts that need be realised in both cases either when migrating or developing from scratch information systems in cloud environments. The main contribution of the paper is the identification of various security and privacy concepts as well as the introduction of new ones under one framework. The identification was based on the threats revealed from the respective literature and extended with the introduction of some newly defined complementary concepts. Future steps include the transformation of these concepts on technical requirements and the design of a modeling process for applying these requirements on a real case scenario.

References

1. Cloud Security Alliance "Top Threats to Cloud Computing V1.0" (retrieved September 22, 2012), https://cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf
2. Heiser, J., Nicolett, M.: Assessing the Security Risks of Cloud Computing, white paper, Gartner group, ID Number: G00157782 (published June 3, 2008)
3. Subashini, S., Kavitha, V.: A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications 34(1), 1–11 (2010)
4. Cloud Security Alliance "Security Guidance for Critical Areas of Focus in Cloud Computing V3.0" (retrieved September 22, 2012), https://cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf
5. Draft, EU Directive for Security issues in Cloud Computing (2012)
6. Kalloniatis, C., Kavakli, E., Gritzalis, S.: PriS Methodology: Incorporating Privacy Requirements into the System Design Process. In: Mylopoulos, J., Spafford, G. (eds.) Proceedings of the SREIS 2005 13th IEEE International Requirements Engineering Conference – Symposium on Requirements Engineering for Information Security. IEEE CPS, Paris (2005)
7. Kalloniatis, C., Kavakli, E., Gritzalis, S.: Addressing privacy requirements in system design: The PriS method. Requirements Engineering 13(3), 241–255 (2008)

8. Kalloniatis, C., Kavakli, E., Kontellis, E.: PRIS tool: A case tool for privacy-oriented Requirements Engineering. Journal of Information Systems Security 6(1), 3–19 (2010)
9. Kavakli, E., Kalloniatis, C., Loucopoulos, P., Gritzalis, S.: Incorporating Privacy Requirements into the System Design Process: The PriS Conceptual Framework. Internet Research, Special issue on Privacy and Anonymity in the Digital Era: Theory, Technologies and Practice 16(2), 140–158 (2006)
10. Kalloniats, C., Kavakli, E., Gritzalis, S.: Dealing with Privacy Issues during the System Design Process. In: 5th IEEE International Symposium on Signal Processing and Information Technology, Athens, Greece, December 18-21 (2005)
11. Mouratidis, H., Kalloniatis, C., Islam, S., Huget, M.P., Gritzalis, S.: Aligning Security and Privacy to support the development of Secure Information Systems. Journal of Universal Computer Science (2012)
12. Mouratidis, H., Giorgini, P.: Secure Tropos: A Security-Oriented Extension Of The Tropos Methodology. International Journal of Software Engineering and Knowledge Engineering (2007)
13. Mouratidis, Giorgini, P.: Security Attack Testing (SAT) - testing the security of information systems at design time. Inf. Syst. 32(8), 1166–1183 (2007)
14. Wei, L., et al.: Managing Security of Virtual Machine Images in a Cloud Environment (2009)
15. Microsoft Technical report: Privacy in the cloud computing era, a Microsoft perspective, Microsoft Corp., Redmond, USA (November 2009)

A Methodology for the Development and Verification of Access Control Systems in Cloud Computing

Antonios Gouglidis and Ioannis Mavridis

Department of Applied Informatics, University of Macedonia
156 Egnatia Str., 54006, Thessaloniki, Greece
{agougl,mavridis}@uom.gr

Abstract. Cloud computing is an emergent technology that has generated significant interest in the marketplace and is forecasted for high growth. Moreover, Cloud computing has a great impact on different type of users from individual consumers and businesses to small and medium size (SMBs) and enterprise businesses. Although there are many benefits to adopting Cloud computing, there are significant barriers to adoption, viz. security and privacy. In this paper, we focus on carefully planning security aspects regarding access control of Cloud computing solutions before implementing them and, furthermore, on ensuring they satisfy particular organizational security requirements. Specifically, we propose a methodology for the development of access control systems. The methodology is capable of utilizing existing security requirements engineering approaches for the definition and evaluation of access control models, and verification of access control systems against organizational security requirements using techniques that are based on formal methods. A proof of concept example is provided that demonstrates the application of the proposed methodology on Cloud computing systems.

Keywords: Security, Inter-organizational systems, Cloud business, Verification.

1 Introduction

Cloud computing is an emergent technology that has generated significant interest in the marketplace and is forecasted for high growth. Specifically, Cloud computing is expected to be a significant growth driver in global IT spending. In fact, by 2013, about 82% of all net-new software firms coming to market will be operationalized on service based software versus a packaged product. By 2016, approximately 25% of all new business software purchases will be of service-enabled software. Software-as-a-Service (SaaS) delivery models will constitute about 14.9% of worldwide software spending across all primary markets and 18% of applications spending, according to IDC [1]. Furthermore, according to the latest research reports, the global computing market is expected to grow at a 30% CAGR reaching $270 billion in 2020 [2]. Therefore, Cloud system should be carefully designed to be secure since otherwise, the existence of security flaws may lead to a decrement in income of an organization.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 88–99, 2013.

In this paper, we are concerned with the development of access control (AC) systems for distributed and collaborative systems, as Cloud systems. Since AC systems are highly complex and of great significance, their right definition, design and development are mandatory for the production of an AC system that corresponds to the initial requirements. To manage the increased complexity of AC systems, system engineering processes are applied. Nevertheless, generic approaches can be enhanced to facilitate the definition of access control requirements in Cloud environments as presented in [3]. Additionally, a common stage in the development of a system is its verification, where a system can be verified for its correctness. Several papers have examined the automated verification of AC systems, and a number of techniques have been proposed to verify them as presented in [4], [5], [6], and [7].

The structure of the remainder of this paper is: Section 2 provides preliminary information regarding AC in the Cloud and systems engineering stages. Section 3 describes our proposed methodology. A use case is given in Section 4. Finally, we conclude this paper in Section 5.

2 Preliminaries

In this section we provide basic preliminary information regarding AC and its role in Cloud systems. Moreover, we briefly refer to two important stages met in any development approach, viz. requirements engineering and verification in the context of AC.

2.1 Access Control

AC in modern distributed systems has become even more challenging since they are complicated and require the collaboration among domains. A domain can be defined as a protected computing environment, consisted of users and resources under a same AC policy. AC is an essential process in all systems. The role of an AC system is to control and limit the actions or operations in a system that are performed by a user on a set of resources. Nevertheless, an AC system is considered of three abstractions of control, namely AC policies, AC models, and AC mechanisms. A policy can be defined as a high-level requirement that specifies how a user may access a specific resource and when. AC policies can be enforced in a system through an AC mechanism that is responsible for permitting or denying a user access upon a resource. An AC model can be defined as an abstract container of a collection of AC mechanism implementations, which are capable of preserving support for the reasoning of the system policies through a conceptual framework. Consequently, the AC model is capable of bridging the existing abstraction gap between the mechanism and the policy in a system [8], [9].

The Cloud is a fairly new and emergent technology and its definition is a topic for discussion in several research papers [10]. Nevertheless, Cloud computing is defined in [11] using five attributes viz. multitenancy, massive scalability, elasticity, pay as you go and self-provisioning of resources. These attributes successfully imprint the distinctive characteristics of the Cloud and differentiate it from similar technologies,

as the Grid computing paradigm. Multitenancy refers to the business model implemented by the Cloud, where a single shared resource can be used from multiple users. Massive scalability refers to the potential of the system to scale (i.e. increase or decrease) in resources. The on-demand and rapid increment or decrement of computing resources is translated as elasticity of the Cloud. Thus, more storage space or bandwidth can be allocated when required, and vice versa. Pay as you go is the process of paying for the resources that are used. Lastly, the users are provided with the ability to self-provision resources, namely storage space, processing power, network resources and so on. An additional characteristic defined in [12] by the National Institute of Standards and Technology (NIST) is Broad Network Access, which states that available capabilities can be accessed using standard mechanisms over the network, and promote their use by heterogeneous clients.

The service model of Cloud computing is based on the SPI framework [11], [12]. SPI stands for Software-as-a-service (SaaS), Platform-as-a-service (PaaS) and Infrastructure-as-a-service (IaaS). Specifically, the SaaS provides software that is used under a business model, namely the usage-based pricing. The PaaS offers the platform for the development of the applications, and lastly, the IaaS handles the provision of the required hardware, software and equipment, in order to deliver a resource usage-based pricing model.

Moreover, the aforementioned service models are provided under three deployment models viz. public, private and hybrid Cloud [11]. The public Cloud provision resources over the Internet and are accessible via a web application. A third-party operates as the host and performs all the required operations (e.g. management, security). The private Cloud provides the same functionality as the public deployment model within internal and private networks. This model requires the acquisition of the appropriate hardware and software. The hybrid model refers to the combination of the public and private deployment models. Usually, the latter model is used to keep sensitive data in the private network and deploy non-core applications to the public. An additional service model proposed by NIST is the community Cloud [12], which refers to infrastructure exclusive used by a specific community of consumers from organizations that have shared concerns.

In Cloud systems, the main objective of AC is to grant authorized users the right to use a service, and at the same time to prevent access to non-authorized users. Similarly to the Grid paradigm, a Cloud AC policy can be defined as a Cloud security requirement that specifies how a user may access a specific resource and when. Such a policy can be enforced in a Cloud system through an AC mechanism, which is enforced by a Cloud Service Provider (CSP). The latter is responsible for granting or denying a user access upon a service. Therefore, AC in Cloud systems is similar to the Grid. The main difference is mostly subject of the service delivery model that is applied on the Cloud system (SaaS, PaaS, IaaS) [11]. Hence, in the SaaS delivery model, the CSP is responsible for managing all aspects of the network, server and application infrastructure. In the PaaS delivery model, the customer is responsible for AC to the applications deployed in the PaaS platform. Lastly, in the IaaS delivery model, the customers are entirely responsible for managing all aspects of AC. In general, AC in the Cloud is not standardized across CSPs, and user AC to Cloud resources is generally weak because of coarse user access management [11].

2.2 Requirements Engineering

This section is intended as an overview of requirements engineering, which is being increasingly recognized as the most critical phase of the systems development process. If the requirements for a system are not right, there will inevitable be problems after the system is delivered. There are immense variations in what is generally understood as a requirement. The term requirement might be used to refer to statements that are clearly at totally different levels of detail. These variations in the statement of requirements arise because of the different ways requirements are used in different organizations. Therefore, a statement of a requirement can range from a high-level abstract statement of a service to be provided or of a system constraint to a detailed mathematical functional specification of a system component [13].

Requirements can be identified into four types: Functional requirements, where define some functionality to be provided by the system (i.e. what the system will do). Non-functional requirements, where define some operational constraints on the behaviour of the system (i.e. how the system should behave). Design requirements, where define constraints on the system design or implementation. Process requirements, where define constraints on the system development process.

Although it is convenient to classify requirements into different classes, there is not really a strict distinction between these. Specifically, non-functional requirements at one level of abstraction are translated into more detailed functional requirements.

Regarding requirements there is a need for specifications at different levels of detail. The principal reason for this is that these different specifications are designed for different purposes and readers. Requirements definition or specifications can be classified into three types viz., user requirements definition, system requirements specification, and software or hardware specifications.

User requirements definition has to be understandable by potential end-users of the system and their managers as well as the developers of the system. The system requirements specification is a more detailed document that is usually part of the contract for the system. Lastly, software or hardware specifications include a detailed description that can serve as a basis for a design or implementation.

Figure 1 illustrates the activities of the requirements engineering process. Despite the fact that there is a clear division between process activities, in practice, the process is iterative with the activities interleaved and with very blurred boundaries between them. Specifically, analysis definitions and specification may be seen as a single activity. In more detail, the feasibility study activity is clearly related to the feasibility study activity that takes place during the conceptual design. Requirement analysis is the process of finding out the requirements and requirement's definitions are intended to be a high-level description of the system requirements. Lastly, system requirement's specifications define the requirements in detail as a basis for the contract for the system procurement and for the system developers.

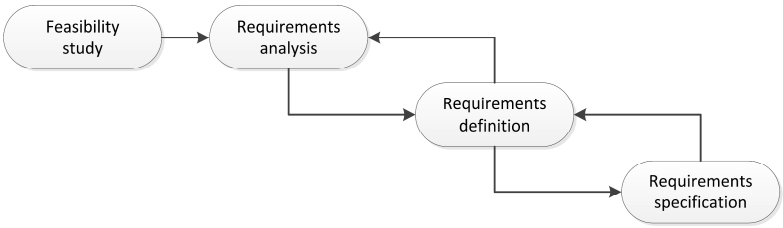


Fig. 1. Requirements engineering process

2.3 System Verification

Verification is a critical process well separated from the previous stages of requirements engineering and system's design, and concerned with building the system right. Verification is used in the comparison of the initial conceptual system based on defined requirements to the computer representation that implements that conception. Specifically, it must ensure that the system does what it should, does it only the way it should and does not do what is should not do [14].

The principal methods for the verification of complex systems are four [15]:

- Testing: Testing is performed on the system itself. However, testing of distributed systems is not always a cost effective process since it can be performed when an implementation of the system is available. Furthermore, it can only prove the existence of bugs, but not their absence.
- Simulation: Simulation-based approaches ensure that a finite number of user-defined system trajectories meet the desired specification. Even though computationally inexpensive simulation is ubiquitous in system design, it suffers from completeness as it is impossible or impractical to test all system trajectories. Furthermore, simulation-based testing is semi-automatic since the user must provide a large number of test cases [16].
- Deductive verification: Deductive verification is based on manual mathematical proof of correctness of a model of a system. It is a very highly cost process and, furthermore, requires highly skilled personnel.
- Model checking: Model checking performs exhaustive testing of all behaviours of a model of the system. It is not vulnerable to the like-hood that an error is exposed; this contrasts with testing and simulation that are aimed at tracing the most probable defects. Additionally, it provides diagnostic information in case a property is invalidated, which is very useful for debugging purposes. In principle, model checking is an automated process and its use requires neither a high degree of user interaction nor complex test data [14].

3 Proposed Methodology

In this section, we provide information regarding our proposed methodology for the development of AC systems. The methodology is independent of the applied development model of a system since the stages of requirements engineering and verification exist in most of them (e.g. sequential, spiral). Figure 2 illustrates the proposed methodology in a system development process that consists of the stages of requirements engineering, systems' design and verification. We propose for the stage of requirements engineering to apply the Conceptual Categorization (CC) [3] and during verification to apply model checking techniques. CC is capable of tailoring the requirements engineering through a re-engineering process. Furthermore, CC operates as an input for the verification stage by defining security requirements. The security requirements are transformed into security properties in temporal logic (e.g. LTL, CTL). The set of defined security properties can be verified on the systems' transition system (TS) using model checking techniques.

In our proposed methodology, we are mostly concerned in performing security requirements engineering and verification of the system to be developed. Therefore, we are not concerned with other stages of the development process, as the design and implementation of the system. However, we depict in Figure 2 their interaction with our defined stages. Thus, the proposed stages can be used transparently in any development model without breaking it.

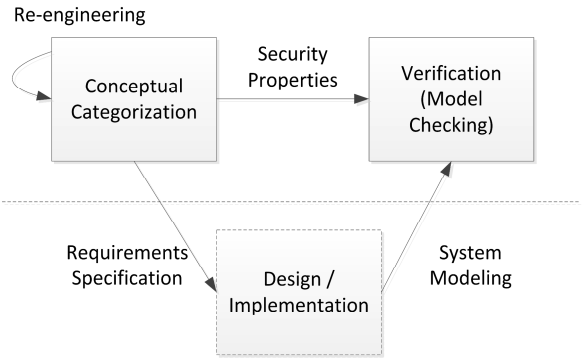


Fig. 2. System development process

3.1 Access Control Requirements

Concerning the requirements engineering stage, we choose to apply the CC that is proposed in [3], which is a layered approach for the definition and evaluation of security requirements for modern collaborative systems, as the Cloud. In brief, CC is based on four abstraction layers. These are the entropy, assets, management and logic layers. The differentiation from generic security engineering approaches is that, in CC, factors that affect the security of the systems are mainly considered in their categorization. CC is able to identify and group security requirements into discrete layers

of different abstraction level. The abstraction level refers to the ability of a layer to identify requirements in different breadth and depth. The entropy layer identifies requirements from the dispersion of the objects in a system and the assets layer from the type of shared objects within the boundaries of the entropy layer. The next layer defines requirements from policy management and the logic layer incorporates requirements that are not handled by the former layers.

In addition to the simple, sequential development process, the CC approach can enhance and be used for tailoring the process of requirements engineering. This is required since in practice new requirements always emerge that inevitably leads to implementation changes. CC can be used for re-engineering existing systems. Therefore, it can be applied on old systems, which usually have a considerable amount of intelligence and experience encapsulated within them. Figure 3 depicts the basic steps that help to re-engineer existing systems using an evolutionary life cycle. The process can be seen as a spiral where cycling through the different processes leads to the desired outcome. Specifically, the evolutionary life cycle includes the definition of requirements in CC layers. In turn, an evaluation of the examined system is performed. This second process, checks the compliance of the already defined requirements. If expectations from the examined systems are low, then new requirements are identified and defined, which when applied will potentially lead to a new system. New requirements are redefined in CC layers and the process is repeated. This evolutionary life cycle helps for faults to be found more quickly and provides the opportunity to include updated technology, and at the same time it facilitates the whole process of delivering a fully functional system [17].

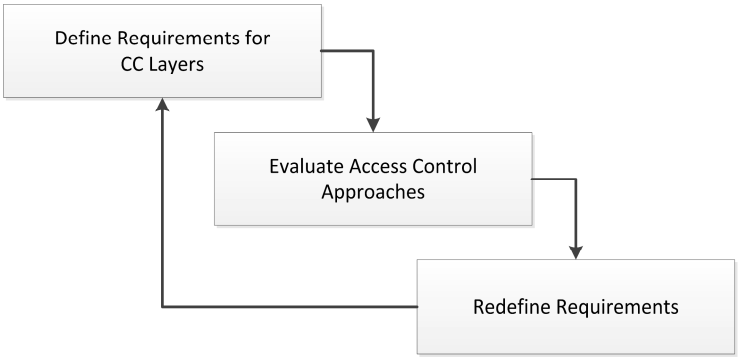


Fig. 3. Re-engineering in Conceptual Categorization

3.2 System Verification

As presented in section 2.3, verification is a process of significant importance. Nevertheless, there are several approaches to perform it, each having its pros and cons. However, in our methodology we choose to apply a model checking technique to verify AC systems for reasons of automation, diagnostic information, integration in existing development cycles and its sound and mathematical underpinning. Model

checking requires the definition of a system's model. These models are characterized as transition systems (TS) and are defined to describe their behaviour.

Formally, a TS is a tuple $(S, Act, \rightarrow, I, AP, L)$ where [14]:

- S is a set of states,
- Act is a set of actions,
- $\rightarrow \subseteq S \times Act \times S$ is a transition relation,
- $I \subseteq S$ is a set of initial states,
- AP is a set of atomic propositions, and
- $L : S \rightarrow 2^{AP}$ is a labelling function.

We consider the TS to be a finite system. TS is called finite if *S*, *Act* and *AP* are finite. Additionally, given that Φ is a propositional logic formula then $s \in S$ satisfies the formula Φ if the evaluation induced by *L(s)* makes the formula Φ true. After defining a TS, the model of the system can be verified against a number of security properties stated in the initial requirements definition stage. The majority of security requirements regarding AC can be expressed as security properties. Security properties are invariants that are given by a condition Φ for the states and requires that Φ holds for all reachable states [14].

4 Use Case

In this section, we demonstrate the proposed methodology by recalling the application of CC on a generic Cloud environment to evaluate and enhance the Role-Based Access Control (RBAC) model [3]. A new produced AC model entitled domRBAC [18] is evaluated in the context of CC. Also, selected requirements were verified in domRBAC using a model checking approach by NIST.

The operational environment is illustrated in Figure 4. The Cloud environment is composed of individually administered domains, which can dynamically join in or quit the collaboration. Users from the participating domains can request on demand usage of services. Specifically, the environment is composed of companies A and B, represented by domains A and B, respectively. An Application Service Provider (ASP) is a corporate organization that can share a number of pay-per-use services. A complementary entity provides a computational computing infrastructure (CCI). Users from companies A and B can request capabilities from their local domain, collaborate with other users, manage their data and request on demand the use of services. The application of CC concludes to the following [3]:

Entropy requirements: The virtual distribution level of the system is low since there is only one formatted organization. The geographic distribution level that depends on the number of the participating domains can be high and entails heterogeneity issues. In order for the AC system to limit access to participating objects, it must be able to successfully authenticate them since domains might make use of different authentication protocols. Furthermore, since the organization's formation is not static, the AC system must continually observe all kinds of modifications.

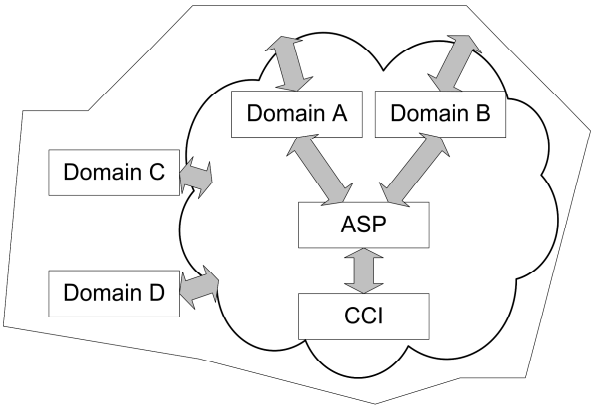


Fig. 4. Operational environment

Assets requirements: AC must be enforced on different types of assets. The scenario considers fine-grained AC on data since it requires sending for computation only segments of users' data. The ASP provides a number of services and the CCI a number of hardware resources. AC for both service and hardware level can be characterized as coarse-grained. The AC system must be able to enforce fine-grained AC on data and coarse-grained on services and hardware resources, respectively.

Management requirements: A number of services uses segments of users' data and submits them at the CCI. This requires a delegation mechanism. Thus, the AC system must be able to support delegation of access rights from domain users to the ASP and CCI. A security issue is that of delegated rights revocation. We assume that delegated rights must be revoked after the completion of a job or on demand by the user. The former requirement demands from the AC system an automation level and the latter to apply changes dynamically. Furthermore, trust relationships must exist between the involving parties. Furthermore, policy conflict resolution must also be examined when composite services exist.

Logic requirements: During user collaboration, their access at the CCI might be restricted by the owner of the resource. This requires an AC system that should support dynamic collaborations. Furthermore, support of stateful sessions because of long lived transactions and decomposition of composed services are required.

Based on the aforementioned requirements, we further proceed with the evaluation of the RBAC model in the context of CC. RBAC handles better centralized architectures where participants are known a priori. Additionally, RBAC is rather coarse-grained approach when it comes to assets definition. Assets, in RBAC, are grouped under roles and to become more granular, the assignments must be split into more. However, the use of context variables in known RBAC variations [19] overcomes such limitations. Delegation of rights and trust relationships are supported, and policy conflict resolution can be tackled in RBAC. Revocation of user assignments, hierarchies and temporal constraints are some of RBAC's. However, RBAC cannot support interactive environments, and also lacks an obligation mechanism, which usually

corresponds to business requirements. Table 1 summarizes the evaluation of RBAC in the context of CC based on [3].

Table 1. Comparison of AC models

AC Model	CC Layers			
	Entropy	Assets	Management	Logic
RBAC	Low / Medium	Low / Medium	Medium / High	Medium
domRBAC	High	Medium	High	Medium/ High

Therefore, after evaluating RBAC, we proceeded with the design of a new access control model entitled domRBAC [18] that is suitable for modern collaborative environments. Table 1 illustrates the evaluation of domRBAC compared to RBAC, where it is shown that domRBAC strengthens the RBAC approach in all layers. Specifically, the features in domRBAC that strengthens RBAC are: i) In entropy layer, domRBAC is capable of supporting secure inter-operation and has the ability to scale well. ii) In assets layer, we identify that domRBAC is able to support basic usage control and therefore usage restrictions can be introduced. Moreover, it is possible to define resources that are being shared by multiple stakeholders. iii) In management layer, using a small administrative overhead it is able to automate the management of policies in an easy and efficient way. iv) In logic layer, domRBAC strengthens RBAC since it provides features such as autonomy and security.

We further refer to the verification technique for the requirements of autonomy and security, which are identified in the logic layer of CC and supported in domRBAC. Secure inter-operation in collaborative systems is required for secure collaboration among participating parties such that the principles of autonomy and security can be guaranteed [20]. The principle of autonomy states that if an access is permitted by an individual system, it must also be permitted under secure inter-operation. The principle of security states that if an access is denied by an individual system, it must also be denied under secure inter-operation. Both principles can be characterized as security properties of a system, which should be preserved during collaborations since their enforcement means that something bad never happens [14].

For the verification of secure inter-operation properties, we applied the technique proposed in [21], which focuses on the verification of generic properties for AC models. The technique is able to cope with various types of AC properties including static, dynamic and historical. It also supports the generation of test cases to check the conformance between models and policy rules through combinatorial test array [22], and optionally generate the verified AC policies in eXtensible Access Control Markup Language (XACML) version 2.0 or 3.0, which is becoming the de facto language for the specification of policy rules in modern collaborative systems such as the Cloud. We adopt the finite state machine to describe the transitions of the authorization states, and the usage of static constraints so to adequately cover the verification of secure inter-operation properties in RBAC. The technique is to verify specified AC properties against AC models using a black-box model checking method [6]. An

implementation -- Access Control Policy Tool (ACPT) [23] is developed by NIST Computer Security Division in corporation of North Carolina State University.

ACPT provides graphical user interface templates for composing AC policies and properties. Checking for conformance of AC properties and models is through the SMV (Symbolic Model Verification) model checker. In addition, ACPT provides a complete test suite generated by NIST's combinatorial testing tool ACTS [22] and an XACML policy output for the verified model. Through these four major functions, ACPT performs syntactic and semantic verifications as well as the interfacing for composing and combining AC policies. ACPT assures the efficiency of specified AC policies, and detects policy faults that leak or prohibit legitimate access privileges.

5 Conclusions

In this paper, we proposed a methodology for the development and verification of AC systems. The development stage is based on CC and the verification stage on a sound and mathematical underpinning technique (i.e. model checking). The latter is feasible via the definition of the system's TS and security properties, which have to be verified. The proposed methodology can be applied on any existing development process since it does not break the development model. Through an example, we demonstrated the proposed methodology, which resulted in a critique of existing AC approaches in [3] and the development of a new AC model for collaborative systems (e.g. Cloud environments) in [18]. Furthermore, we referred to the application of a model checking approach proposed by NIST, which resulted in the verification of the principles of security and autonomy in our defined AC model.

Acknowledgement. This work has been (partially) funded by the Research Committee of the University of Macedonia, Greece.

References

1. Mahowald, R.P., Sullivan, C.G., Konary, A.: Market Analysis Perspective: Worldwide SaaS and Cloud Services, 2012 — New Models for Delivering Software (2012), http://www.idc.com/getdoc.jsp?containerId=238635#.UM4QauTqmuM
2. Media, M.R. Global Cloud Computing Market Forecast 2015-2020 (2012), http://www.marketresearchmedia.com/?p=839
3. Gouglidis, A., Mavridis, I.: On the definition of access control requirements for grid and cloud computing systems. In: Doulamis, A., Mambretti, J., Tomkos, I., Varvarigou, T. (eds.) GridNets 2009. LNICST, vol. 25, pp. 19–26. Springer, Heidelberg (2010)
4. Hansen, F., Oleshchuk, V.A.: Conformance checking of RBAC policy and its implementation. In: Deng, R.H., Bao, F., Pang, H., Zhou, J. (eds.) ISPEC 2005. LNCS, vol. 3439, pp. 144–155. Springer, Heidelberg (2005)
5. Jayaraman, K., et al.: Automatic error finding in access-control policies. In: Proceedings of the 18th ACM Conference on Computer and Communications Security. ACM (2011)

6. Vincent, C., et al.: Model checking for verification of mandatory access control models and properties. International Journal of Software Engineering and Knowledge Engineering 21(01), 103–127 (2011)
7. Fisler, K., et al.: Verification and change-impact analysis of access-control policies. In: Proceedings of the 27th International Conference on Software Engineering. ACM (2005)
8. Capitani di Vimercati, S., Foresti, S., Samarati, P.: Authorization and Access Control. In: Petković, M., Jonker, W. (eds.) Security, Privacy, and Trust in Modern Data Management, pp. 39–53. Springer, Heidelberg (2007)
9. Sandhu, R.S., Samarati, P.: Access Control: Principles and Practice. IEEE Communications Magazine 32(9), 40–49 (1994)
10. Foster, I., et al.: Cloud computing and grid computing 360-degree compared. In: Grid Computing Environments Workshop, GCE 2008. IEEE (2008)
11. Mather, T., Kumaraswamy, S., Latif, S.: Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance. O'Reilly Media, Inc. 292 (2009)
12. Mell, P., Grance, T.: The NIST definition of cloud computing (draft), vol. 800, p. 145. NIST Special Publication (2011)
13. Sommerville, I., Kotonya, G.: Requirements engineering: processes and techniques. John Wiley & Sons, Inc. (1998)
14. Baier, C., Katoen, J.P.: Principles of model checking, vol. 26202649. MIT Press (2008)
15. Heljanko, K.: Model Checking based Software Verification (2006), http://iplu.vtt.fi/digitalo/modelchecking.pdf
16. Girard, A., Pappas, G.: Verification using simulation. Hybrid Systems: Computation and Control, 272–286 (2006)
17. Stevens, R.: Systems engineering: coping with complexity. Pearson Education (1998)
18. Gouglidis, A., Mavridis, I.: domRBAC: An access control model for modern collaborative systems. Computers & Security (2012)
19. Tolone, W., et al.: Access control in collaborative systems. ACM Computing Surveys (CSUR) 37(1), 29–41 (2005)
20. Gong, L., Qian, X.: Computational issues in secure interoperation. IEEE Transactions on Software Engineering 22(1), 43–52 (1996)
21. Hu, V.C., Kuhn, D.R., Xie, T.: Property verification for generic access control models. In: IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, EUC 2008. IEEE (2008)
22. Kuhn, R., Lei, Y., Kacker, R.: Practical combinatorial testing: Beyond pairwise. IT Professional 10(3), 19–23 (2008)
23. Hwang, J.H., et al.: ACPT: A tool for modeling and verifying access control policies. In: 2010 IEEE International Symposium on Policies for Distributed Systems and Networks (2010)

Simple Mail Delivery Protocol

Recipient-Based Email Delivery with Anti-spam Support

Dimitris Zisiadis, Spyros Kopsidas, and Leandros Tassiulas

University of Thessaly and Centre for Research & Technology Hellas, Volos, Greece
{dimitris,spyros,leandros}@uth.gr

Abstract. In this paper we propose a user based architecture for the email system, where the recipient of an email message manage its receipt in conjunction with both origin and destination email servers. Messages are kept to the origin email server until a confirmation from the recipient is issued though the destination email server. Therefore, spam email doesn't travel the Internet and doesn't overload the destination email server and recipient's mailbox. White and black lists for (sender, email_server) pairs are built automatically following user evaluation. With our proposal, full control is delegated to the recipient over the email activity. On the server side, Simple Mail Delivery Protocol (SMDP) manages email delivery and options. SMDP server acts as a gateway for sending emails, a repository for the users' mailboxes and a central point for managing user options related to email filtering and spam handling. The solution is suitable for group/business email handling.

Keywords: e-mail, privacy; anti-spam; SMDP, SMTP.

1 Introduction

Since the dawn of the Internet, email was one of the most important applications and since then it has emerged as a primary mode of communication, especially in the business sector. It has largely replaced traditional means of communications, like fax and paper mail. It has even eliminated the need to have telephone calls in many cases. Email is the Internet counterpart to the paper mail; everybody is free to use the email system to send emails to valid email addresses. The email system was designed originally as the counterpart of the paper mail; anyone can send email to a valid email address, without any involvement of the recipient who retrieves any emails from his local mailbox.

Spam is a major problem of email systems nowadays. Spam is the use of the email system to send unsolicited bulk messages indiscriminately. Typical anti-spam techniques are used in the destination domain, after message receipt, either at a server level or at the user level. It is self-evident for the networked employees that the volume of received email messages is huge, while spam email volume is overwhelming. Spamming is both annoying and usually unethical and despite all the efforts to eliminate it, the amount of spam email compensates constantly for more

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 100–111, 2013.

than 80 percent of the total email volumes [8]. In response to spam, spam-filters have moved from simple repositories to dynamic knowledge locates on local servers. In spite of the development of strong defense mechanisms, users still receive huge amounts of spam emails into their mailboxes. Moreover, filtering mechanisms in their effort to identify as much spam emails as possible, mark legitimate emails, matching some of the criteria for spam, as false-positive spam.

In this paper we propose a new architecture for the email system for group/business emails, where users have full control over their email activity over the Internet; the origin email server is in charge of accepting emails from registered users only and delivering them in coordination with the recipients of the message through the destination email servers. We propose a new protocol for email delivery, the Simple Mail Delivery Protocol (SMDP), to replace SMTP for group/business email activity, in order to enhance user confidence and privacy for business email exchange while relaxing the network and the receiving email servers from unnecessary spam email burden. Emails are not allowed to travel over the Internet to the destination email server unless the recipient acknowledges receipt. Every email on the origin email server has to be authorized by its recipient in order to be delivered to the destination email server, whereas spam email is blocked at the origin email server and does not overload the network or the remote email server. Email servers compile white and black lists for their registered users on the fly, by recording user preferences while managing their email volumes. Spammers are identified by the perspective of the email community based on the volume of emails classified as spam for every user of the system.

This paper is organized as follows: an overview of the email system is provided in section 2, while section 3 outlines widely accepted anti-spam techniques. Related work is presented in section 4 while in section 5 we present our proposal for the new redesigned email system; section 6 presents performance issues for our protocol. Finally concluding remarks are given in section 7.

2 E-mail System

Email is the Internet counterpart of the paper mail service, based around the notion of mailboxes for users that are registered to a specific Internet domain. When a user sends an email, the origin server (origin) forwards the message over the network to the recipient's email server (destination). More precisely, there is a Mail Transport Agent (MTA) serving each domain that is used for email message delivery over the Internet. When an email is sent, the origin domain's MTA receives the message from the sender and forwards it to the recipient's domain MTA that in turn delivers the message to the local Mail Delivery Agent (MDA) where the recipient's mailbox is located. The recipient interacts with the local MDA in order to manage its mailbox through the Mail User Agent (MUA). MUA has two forms; it can either be software installed on the user's host in which case it is called an "email client" or a web interface can be used for the same purpose in which case it is called "webmail". Mailbox privacy is ensured through the protection of the MUA to MDA communication via {username, password} authentication scheme. A similar scheme can be used in the origin domain for MUA to MTA communication in order to verify sender's identity.

Simple Mail Transfer Protocol (SMTP) [1-2] is the protocol used for MTA to MTA communication. The most common SMTP sever is Sendmail [3], initially distributed as part of the UNIX [4] operating system. HELO/EHLO command starts an SMTP session that ends with the QUIT command. An email transaction begins with the MAIL command (multiple MAIL commands allowed within the same SMTP session). Recipients are added through the use of the RCPT command. The DATA command starts email message body while a "." on a line by itself ends the respective data entry; header lines (i.e. Subject, Cc, Reply-To, etc.) are also included within the message body. SMTP server listens on well-known port 25. The two most common protocols for retrieving emails from the local MDA are the Post Office Protocol (POP3) [5-6] and the Internet Message Access Protocol (IMAP) [7]. MDAs are called POP servers or IMAP servers, depending on which protocol they use.

POP3 is an open Internet standard and the most common email client connection protocol. POP3 protocol enables email clients, independent of user location, to connect to any email server through a {username, password} authentication scheme, in order to manage user's email activity, performing all the necessary email management functions, i.e. read, send, reply, forward, etc. A POP3 server listens on well-known port 110. Encrypted communication for POP3 is either requested after protocol initiation, using the STLS command, if supported, or by POP3S, which connects to the server using Transport Layer Security (TLS) or Secure Sockets Layer (SSL) on well-known TCP port 995 (e.g. Google Gmail). IMAP is more feature-rich than POP3, supporting both on-line and off-line modes of operation, easing email management for those using more than one device and people on the move, offering the ability to select the emails to be viewed through the list of incoming messages. An IMAP server listens on well-known port 143.

The complete procedure for the email exchange is as follows:

1. *Composition:* Sender composes new message in the email client.
2. *Upload:* Message is uploaded to the SMTP server.
3. *DNS:* SMTP server uses DNS to retrieve MX record for recipient's domain and get the address for the destination SMTP server.
4. *Transfer:* Message is routed through the Internet to the recipient's SMTP server.
5. *Delivery:* Recipient's SMTP server delivers the message to the user mailbox on local POP/IMAP server.
6. *Read:* When recipient gets online (if not already), his email client connects to the local POP/IMAP server and downloads message.

3 The SPAM Problem

Filtering of received emails is the basis for most of the anti-spam policies. The main goal is to identify unsolicited/bulk email and prevent the end user from having to go through it. Hereunder we present briefly the most common approaches. Filtering can be performed either at the server-level or the user-level. Server-level filters set rules for all registered users in the specific domain, whereas user-level filters identify spam when it reaches the user's terminal.

White lists [17] are lists of individual e-mail addresses, IP addresses or domain names considered to be safe. Email address spoofing [18], zombies [19] and botnets [20] make them less efficient. Black lists [21] on the other hand contain addresses that are considered as spam sources. Aggressive black lists may block whole domains or ISP's. IP reputation [22] is an email reputation method applied prior to a message being accepted, where IP reputation lists of legitimate domains and spam domains are maintained. Another such method is content-based signatures [23], which apply after a message has been accepted, where black and white lists are used. When a new message arrives, it is inspected and classified as spam or legitimate. Sender's reputation decreases for every spam email, while it increases when a legitimate email is received. The sender's address is moved to the white or black list depending on a specified reputation threshold. CAPTCHAs [24-25] are techniques that discourage spammers in their efforts to use zombies and botnets, although their use could annoy or delay legitimate users. They force the input of on-screen displayed non-machine-readable data, exploiting the human attributes of email entities in order to prove that the sender is in fact a human. The most common approach is to request from the user to input a string hidden in a picture.

Content-based filtering [26] techniques analyze received email content to examine whether it is legitimate. There are two categories of content based filtering: heuristic filtering and machine learning filtering. Heuristic filters [23] are sets of hand written rules; they can effectively investigate the whole email content or specific parts. The two main disadvantages of heuristic filtering are a) the complexity of rule statements being a problem for the average user and b) they keep spammers motivated to react and invent new ways to avoid detection. Machine learning filters [23] use algorithms that feature the ability to learn from the incoming email activity and increase their efficiency along the way. Bayesian filtering [27-28] is a widely adopted approach with great efficiency. Bayesian filters are "trained" to recognize spam. A well-trained Bayesian filter could achieve a high accuracy rate, over 95% and has the ability to evolve as spam evolves. Finally, collaborative filtering [23] is an anti spam approach in which groups of users cooperate by using the same technique to eliminate spam. P2P-Based Collaborative Spam Detection and Filtering [29] is a widely used approach of the kind.

4 Related Work

Research focus on email protocols usually addresses authentication, privacy, security, spam, fairness and non-repudiability issues. Fairness refers to the assurance that both sides will get the expected messages or neither will; non-repudiability meaning that when an email is successfully sent, the sender cannot deny sending it and the recipient cannot deny receiving it. Below we briefly present the most indicative proposals.

A protocol using an effective RSA-based convertible signature with non-interactive partial signature proof method and additively split signing secret key is proposed in [9], proved to suffer from key exposure from the signer's registration information alone; a newer proposal alleviating this flaw was presented in [10]. Reducing Trusted Third

Party's (TTP) storing demands by using Key Chains, offering strong fairness, non-repudiability and timeliness is proposed in [11], further enhanced in [12], where TTP transparency is added and a weakness of the original protocol is fixed. A proposal ensuring secure copies of email messages on backup systems and intermediaries, especially important for multi-hop transmissions through multiple routers and mail servers is presented in [13]. A password-based authentication protocol adopting signcryption is proposed in [14] guaranteeing resistance to the sender server's forgery attack, confidentiality, forward-secrecy, authentication and non-repudiability. An identity-based authenticated protocol ensuring perfect forward secrecy, authentication, confidentiality and low computation cost is proposed in [15]. The spam problem has also been extensively researched; [16] presents the adopted approaches and proposes a solution based on social networks.

5 The SMDP Protocol

Simple Mail Delivery Protocol (SMDP) is a user-based approach that enables users to have full control over their email activity while addressing the problem of spam in a direct and personalized way, a key requirement for every business activity.

5.1 SMDP's Basic Assumptions

The basic assumptions for our proposal are the following:

1. *Registration*: Every email address tightly couples a user to an SMDP server. The user's email address is the main user identity for that user. To enhance user convenience, additional email addresses can be coupled to the same user through the main user identity. These are called secondary identities.
2. *Connection*: SMDP servers accept connections from main and secondary user identities alone.
3. *No spoofing*: The "From" field of every email message must match either a main user identity or any secondary identity. No email address spoofing should be allowed. In case a server is compromised this may not hold true; compromisation could be detected though server administration policy: the administrator of the compromised server can act based on feedback from its peers.
4. *Message at origin*: The origin email server holds email messages until the intended message recipient issues an acknowledgement for delivery.
5. *Message acceptance*: Upon acceptance of message delivery by its intended recipient, the origin email server forwards the message to the destination domain. When multiple recipients are at the same destination domain, a copy of the message is kept to the destination server after first message acceptance for local delivery.
6. *Message denial*: When the intended recipient rejects delivery of a message, a negative acknowledgement is sent to the origin email server and the message is deleted in its origin.

5.2 Basic SMDP Operation

We present the basic operation of SMDP with an example, where user_A@domain_A sends an email to user_B@domain_B. Let SMDP_A is the SMDP server for domain_A and SMDP_B is the SMDP server for domain_B. Message delivery goes through the following steps, depicted graphically in Fig. 1:

1. *A's Connection*: user_A connects to SMDP_A.
2. *Composition*: user_A composes new message.
3. *Acceptance at origin*: SMDP_A accepts user_A message for delivery.
4. *Email envelope*: SMDP_A sends an "email envelope" to SMDP_B where the envelope is stored in user_B's mailbox. The email envelope contains sender, intended recipient(s), subject, attachment status and a couple of headlines.
5. *B's Connection*: User_B connects to SMDP_B.
6. *Envelope forwarding*: SMDP_B sends email envelope to user_B's email viewer. User_B may select to either accept or reject the email message. SMDP_B receives selection from user_B for this envelope.
7. *Message accepted*:
   (a) SMDP_B requests entire message from SMDP_A.
   (b) SMDP_A sends message to SMDP_B.; SMDP_B stores message in user_B's mailbox; if user is still signed in it is displayed in his viewer.

7'. *Message rejected*:
   (a) SMDP_B sends "reject" for this envelope to SMDP_A.
   (b) SMDP_A deletes message from its queue.

5.3 SPAM Handling

The basic SMDP operation handles email delivery depending on user acknowledgement from the intended recipient of the email message. However, SMDP can also play an active role in spam handling based on user evaluation of the incoming messages.

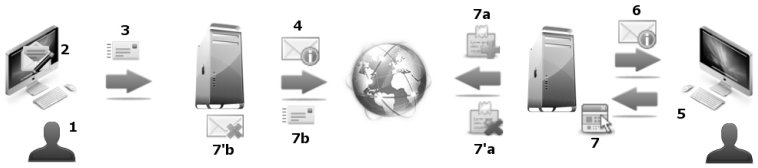


Fig. 1. SMDP operation

This can be achieved by refining recipient options to a more sophisticated scheme than native message accept/reject, while keeping evaluation straightforward and immensely profitable for the user. The available message delivery options are depicted

in Table 1. Note that the "reject" family of user options is available even for previously acknowledged messages, to keep user relationships up to date and to enable user reaction to abnormal spoofing email activity from trusted contacts. This way SMDP gathers personal user evaluation for the incoming messages as well as for sender's reputation. SMDP can take advantage of this personal evaluation history, thus automating a personalized message service from the email system to the specific user.

Table 1. User Options

Option	Meaning
Accept	Accept message.
Accept & authorize	Accept message and authorize sender permanently. Add (sender, SMDP_sender) pair to user's white list.
Reject	Reject message.
Reject & spam	Reject message and mark it as spam.
Reject & spammer	Reject message and block sender as spammer. Add (sender, SMDP_sender) pair to user's black list.

White and black lists are composed on the fly according to email envelopes evaluation. Each SMDP server maintains white and black lists per user, in the form of (email, SMDP) pairs. We do so in order to identify unlawful SMDP servers and/or users (i.e. an unlawful SMDP server can send fraudulent emails, appearing to be sent by users not belonging to its domain or an SMDP server operates as an open relay). These white and black lists operate as filters for received email. In our example, in the case that user_B selects to "accept&authorize" the email received from user_A, the pair (user_A, SMDP_A) is added to user_B's white list on SMDP_B, while if "reject&spammer" is selected, the pair (user_A, SMDP_A) is recorded in user_B's black list on SMDP_B. The SMDP server of the recipient communicates user evaluation data to the origin SMDP server, enabling composition of white and black lists for the respective sender user identity in the origin domain; thus filtering is enabled for the outgoing email activity of the sender of the message in transit. Outgoing white and black lists need to contain only email addresses, as they do not need to keep record of destination SMDP, which is natively bonded to the recipient domain. In our example, (user_B) is recorded to user_A's white list on SMDP_A when user_B selects "accept&authorize" or to user_A's black list on SMDP_A when user_B selects "reject&spammer". The SMDP messages, in accordance to the above user options along with the relevant SMDP operations in the destination and origin domains are defined in Table 2. User lists speed up processing of messages exploiting past email transactions. When the recipient is in the sender's white list (and consequently the sender is in the recipient's SMDP white list) the entire message can be safely sent out without having to go through the envelope phase. On the other hand when the recipient is in the sender's black list the message is rejected immediately at its source. Table 3 summarizes the use of white and black lists in SMDP.

Table 2. SMDP Messages

Message	Recipient SMDP	Origin SMDP
ACCEPT	Send "ACCEPT" to origin SMDP for this envelope.	Send entire message.
AUTHORIZE	Send "AUTHORIZE" to origin SMDP for this envelope; add (sender, SMDP) to the user's white list.	Add recipient to the sender's white list; send entire message.
REJECT	Send "REJECT" to origin SMDP for this envelope.	Delete message.
SPAM	Send "SPAM" to origin SMDP for this envelope.	Investigate with sender; delete message.
SPAMMER	Send "SPAMMER" to origin SMDP for this envelope; add (sender, SMDP) to the user's black list.	Add recipient to the user's black list.

Table 3. Use of White and Black Lists in SMDP

Peer side	*List Entry*	*Meaning*
Sender	Recipient in sender's white list (As a result of sender in recipient's white list).	Authorized user: send entire message directly (no envelope phase).
Sender	Recipient in sender's black list.	Reject message (nothing sent).
Recipient	Sender in recipient's white list.	Authorized user; accept message.
Recipient	Sender in recipient's black list.	Reject message; send "SPAMMER" to the origin SMDP server.

In addition to the above, in every domain, a quality indicator per peer SMDP_X, namely Quality (X), recording reputation of SMDP_X for that domain can be constructed as follows. A local counter, Messages (X), keeps the total number of emails received from this server. Every time an email is received from SMDP_X, the counter is increased. Another counter, Votes (X), keeps the quality score for the messages received from SMDP_X, following recipient user perspective from the local users; this counter is increased for every “accept” while it is decreased for every “reject”. The quality indicator for SMDP_X is defined as the ratio of the two:

$$\text{Quality (X)} = \text{Votes (X)} / \text{Messages (X)}$$

The values of the indicator fall in the range of [-1, 1], being proportional to the quality of the emails reaching recipient domain from all users of SMDP_X. Thresholds can be defined to accommodate for automate management alerts, meeting the local management criteria. For example an SMDP with a negative indicator (more spam than legitimate emails) could result in blocking this SMDP. We have defined reference commands to implement all the aspects of the SMDP protocol as described above. Both server-to-server communication as well as user-to-server communication needs to be formulated. Table 4 presents the implementation design for SMDP peer communications whereas in Table 5 the user application to SMDP server communication is exploited.

Table 4. SMDP Server-to-Server Commands

List Entry	Meaning
HELO/EHLO sendinghostname	Initiates the SMDP session. Multiple envelopes and mails can be sent in the same session.
ENVELOPE From: <source email address>	Indicates the start of an envelope message.
MAIL From: <source email address>	Indicates the start of an email message.
RCPT To: <destination email address>	Recipient of the email, one per recipient.
SIZE=numberofbytes	The size of the message in bytes.
DATA	Start of the email message body, to be terminated by a "." on a line of its own. Email header lines (Subject, Cc, Reply-To, etc.) are sent within the message body.
QUIT	Terminates the SMDP connection.
ACCEPT, AUTHORIZE, REJECT, SPAM, SPAMMER	As described in Table II.

Table 5. User application to SMDP server commands

List Entry	Meaning
user (username)	Login username; if valid, server will request for password.
pwd_ (password)	Send password. If authentication is successful, server responds with the envelopes and messages in user mailbox.
mail	Create new mail (From: field is added automatically by the SMDP server, no spoofing allowed).
rcpt to	Add recipient.
data	Email message body, including any header lines.
list	Get list of messages.
retr (message)	Get message number "message".
del (message)	Delete "message".
accept (envelope)	Request message corresponding to the "envelope".
accept&auth (envelope)	Request message corresponding to the "envelope" and add sender to whitelist.
reject (envelope)	Reject message corresponding to the "envelope".
reject&spam (envelope)	Reject message corresponding to the "envelope" and mark it as spam.
reject&spammer (envelope)	Reject message corresponding to the "envelope" and add sender to blacklist.
quit ()	Quit session.

5.4 Advantages over SMTP

SMDP, while still remaining simple and intuitive much like SMTP, supersedes SMTP in the following:

1. *User control:* SMDP is user oriented whereas SMTP is a network service beyond any form of user control.

2. *Privacy:* SMTP does not address privacy. SMDP, on the other hand empowers users to control their mail activity, providing simple yet effective mechanisms that enhance user privacy. This is achieved by blacklisting undesirable contacts.
3. *User-friendly envelopes:* the envelope format is very common to virtually any Internet user, as it resembles web engine search results.
4. *Communications management:* personal filters are maintained on the SMDP server, enabling users to manage email communications simple and effectively.
5. *Dynamic filtering:* email-filtering options are monitored on the fly, recording user preferences while reading their emails.
6. *Quality indicator*: email server evaluation is automatically constructed per peer SMDP, based on local domain user perspective, providing a valuable indicator to the server administration.

6 Performance Issues

As already mentioned spam emails account for more than 80% of the total email volume. Consequently, a lot of valuable computational and network resources are currently wasted during the handling procedures for all those spam messages. SMDP eliminates the need to transfer over the Internet and process at the destination domain huge volumes of unsolicited emails.

SMDP though requires some additional processing for email envelopes and email messages as well as white and black list maintenance. For the calculation of the overheads imposed by the protocol during its operation we used a typical mail server hardware setup: 2.13 GHz Intel Core2Duo (E6400) 64-bit CPU and 2GBs of RAM, running Ubuntu Linux 9.1 32-bit. We measured the basic SMDP time overheads for the following operations:

1. *Email address match*: Search for an email address into a text file. If the email address is not matched it is inserted into the file. This procedure is used in list maintenance, when an email address must be added or removed for a white or black list. The procedure was executed using 20 different email addresses in a 1000 lines text file along with ten “add” commands and another ten “delete” commands. The average execution time for each combined search and add/delete operation was 9msec.
2. *Email envelope extraction*: This procedure is used every time an email is sent from an unauthorized user. We used 281 different email messages with a total size of 191MBs. In order to achieve real world conditions, we restricted the maximum CPU resources available to that procedure to 37%. The overall time required for processing the total email volume was 14.391sec, i.e. an average time of 51msec for each message.

The performance results of the operations enhance the aspect that applying the SMDP protocol into the core email communications does not produce significant overheads to the fundamental system resources. As far as disk usage is concerned, extra disk space is needed at the origin email servers to keep the envelopes until the mail is

either forwarded or rejected by the recipient; on the other hand less disk space is used at the destination server where spam mail volume is reduced to the envelope size only instead of the full email message.

7 Conclusions and Future Work

In this work we presented the basic principles of Simple Mail Delivery Protocol, a novel protocol for email exchange as an alternative to the traditional SMTP. SMDP is a user based email message receipt protocol that can handle email exchange much like SMTP; moreover, SMDP supports spam handling natively, maintaining personalized filters on the fly and recording user preferences as they manage their emails. Inbound white and black lists per user are maintained locally for receiving emails. Outbound white and black lists are also compiled per user as a result of user's emails sent over the Internet. SMDP establishes trust relationships between email users for which email delivery is automated and pre approved. It also keeps unwanted email messages from being transferred over the network and overloading user mailboxes, while it makes spammer identification based on user opinion. Moreover, a local domain indicator reflecting the quality for every SMDP peer is automatically calculated based on the total emails received from the peer SMDP users, providing domain operators a valuable management tool.

In this paper we presented the principle of operation for our protocol. For the future we plan to develop the software to implement SMDP and operate and evaluate a pilot based on our design.

Acknowledgments. This work is part of the project "Network of Excellence in Internet Science" ICT-FP7- 288021 NoE EINS, funded by the European Commission.

References

1. Klensin, J.: Simple Mail Transfer Protocol. RFC5321 (2008)
2. Rose, M.: SMTP Service Extensions. RFC 1425 (1993)
3. Sendmail, `http://www.sendmail.com/sm/open_source/`
4. Ritchie, D., Thompson, K.: The UNIX Time-Sharing System. The Bell System Technical Journal (July-August 1978)
5. Myers, J., Rose, M.: Post Office Protocol – Version 3. RFC 1939 (1996)
6. Gellens, R., Newman, C., Lundblade, L.: POP3 Extension Mechanism. RFC 2449 (1998)
7. Crispin, M.: Internet Message Access Protocol – Version 4rev1. RFC 3501 (2003)
8. Messaging Anti-Abuse Working Group (MAAWG), `http://www.maawg.org/email_metrics_report`
9. Park, J., Chong, E., Siegel, H.: Constructing fair-exchange protocols for E-commerce via distributed computation of RSA signatures. In: Principles of Distributed Computing, PODC 2003, pp. 172–181 (2003)
10. Wang, H., Ou, Y., Ling, J., Xu, X., Guo, H.: A New Certified Email Protocol. In: 18th International Workshop on Database and Expert Systems Applications, pp. 683–687 (2007)

11. Cederquist, J., Dashti, M.T., Mauw, S.: A Certified Email Protocol Using Key Chains. In: 21st International Conference on Advanced Information Networking and Applications Workshops, pp. 525–530 (2007)
12. Liu, Z., Pang, J., Zhang, C.: Extending a Key-Chain Based Certified Email Protocol with Transparent TTP. In: 2010 IEEE/IFIP 8th International Conference on Embedded and Ubiquitous Computing, pp. 630–636 (2010)
13. Jang, J., Nepal, S., Zic, J.: Trusted Email protocol: Dealing with privacy concerns from malicious email intermediaries. In: 8th IEEE International Conference on Computer and Information Technology, pp. 402–408 (2008)
14. Zhang, J., Chen, H.: An improved password-based authenticated email protocol. In: The 2nd IEEE International Conference on Information Management and Engineering, pp. 545–549 (2010)
15. Zhang, J., Chen, H.: An Efficient Identity-Based Authenticated Email Protocol with Perfect Forward Secrecy. In: 2010 International Forum on Information Technology and Applications, pp. 68–71 (2010)
16. Zisiadis, D., Kopsidas, S., Varalis, A., Tassiulas, L.: Mailbook: A social network against spamming. In: 2011 International Conference for Internet Technology and Secured Transactions, pp. 245–249 (2011)
17. Wikipedia, http://en.wikipedia.org/wiki/Whitelist#Email_whitelists
18. Mathew, A.R., Al Hajj, A., Al Ruqeishi, K.: Cyber crimes: Threats and protection. In: 2010 International Conference on Networking and Information Technology, pp. 16–18 (2010)
19. Zhenhai, D., Peng, C., Sanchez, F., Yingfei, D., Stephenson, M., Barker, J.: Detecting Spam Zombies by Monitoring Outgoing Messages. In: The 28th Conference on Computer Communications, pp. 1764–1772 (2009)
20. Hachem, N., Ben Mustapha, Y., Granadillo, G.G., Debar, H.: Botnets: Lifecycle and Taxonomy. In: 2011 Conference on Network and Information Systems Security, pp. 1–8 (2011)
21. Jung, J., Sit, E.: An Empirical Study of Spam Traffic and the Use of DNS Black Lists. In: 4th ACM SIGCOMM Conference on Internet Measurement, pp. 370–375 (2004)
22. Esquivel, H., Akella, A., Mori, T.: On the Effectiveness of IP Reputation for Spam Filtering. In: 2nd International Conference on Communication Systems and Networks, pp. 1–10 (2010)
23. Sanz, E.P., Hidalgo, J.M.G., Pérez, J.C.C.: Email Spam Filtering. In: Advances in Computers, vol. 74, pp. 45–114 (2008)
24. Wikipedia, http://en.wikipedia.org/wiki/CAPTCHA
25. The Official CAPTCHA site, http://www.captcha.net
26. Wikipedia, http://en.wikipedia.org/wiki/Content_filtering
27. Obied, A.: http://ahmed.obied.net/research/papers/spam_paper.pdf
28. Sahami, M., Dumais, S., Heckerman, D., Horvitz, E.: A Bayesian approach to filtering junk e-mail. In: AAAI Workshop on Learning for Text Categorization, pp. 55–62 (1998)
29. Damiani, E., De Capitani di Vimercati, S., Paraboschi, S., Samarati, P.: P2P-Based Collaborative Spam Detection and Filtering. In: 4th Int. Conference on Peer-to-Peer Computing, pp. 176–183 (2004)

Why Are Users of Social Media Inclined to Word-of-Mouth?

Patrick Mikalef[1], Adamantia Pateli[1], and Michail N. Giannakos[2]

[1] Ionian University, Corfu, Greece
{mikalef,pateli}@ionio.gr
[2] Norwegian University of Science and Technology (NTNU), Trondheim, Norway
michail.giannakos@idi.ntnu.no

Abstract. Motivated by the fact that social media are continuously gaining in popularity, firms are piloting different approaches of promoting their products and services. However, there is much debate in the academic and business community about the effectiveness of social media as a platform for marketing. Specifically, practitioners are concerned with how Word-of-Mouth (WOM) is spread through these sites, and what aspects facilitate users in doing so. The aim of this research is to elucidate the motivations for WOM over social media based on utilitarian and hedonic theories. Partial Least Squares (PLS) analysis is performed on data obtained from 169 social media users to test our research model. Outcomes indicate that both, utilitarian and hedonic motivations impact WOM. Utilitarian motives are triggered through customized advertisements, while hedonic motives through the ability to socialize during product browsing.

Keywords: Word of Mouth, Social Media, Utilitarian and Hedonic Theory.

1 Introduction

Word-of-Mouth (WOM) has been defined as the mechanism by which humans exchange information and experiences about a wide range of topics, such as products, companies and services [10]. WOM is recognized as one of the most reliable sources of information transmission since the begging of human society, and has been proven to greatly affect consumers buying decisions [27]. Consequently, it has managed to attract the attention of many researchers, in the attempt to understand what motivates consumers to pass on information. However, with the inception of the Internet, and the bidirectional communication capabilities which it facilitates, WOM has been shifted from an interpersonal to a global scale. As a result, WOM transmitted online has dominated the attention of marketers, since the type of information exchanged vastly impacts consumer purchasing decisions.

Research studies estimate that while 90% of WOM conversations take place offline, the information that is exchanged in these conversations is largely based on opinions found on the Internet [4]. Therefore, consumer motivations to share their positive or negative feedback is of great importance for the sales of companies. Interpersonal communication theories have tried to demystify the tendencies of consumers

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 112–123, 2013.

to pass on their opinions. However, they have been rendered as inappropriate when applied to online WOM, since other motives may propel consumer actions. Consequently, efforts have recently been focused on explaining the psychological drivers for online WOM. While traditionally researchers have focused on aspects of products/services and antecedents that drive users to pass on information, recent attempts have recognized the importance of the medium in facilitating WOM [10].

Online WOM has seen a great rise, with Amazon.com being one of the first companies to introduce the feature of opinion giving on purchased products. Today, it has an estimated 10 million consumer generated comments [5] Following the successful example of Amazon.com, numerous companies introduced this feature allowing their customers to provide feedback on purchased items. Additionally, websites specialised on opinion sharing have risen, such as epinions.com and buzzilions.com. These sites provide a concentrated source of feedback for products of numerous types. Recently, there has been an increasing amount of attention been placed on marketing over social media websites as an alternative option for product promotion and customer feedback [24]. Due to their large user base, companies have piloted attempts to harness the potential that such sites may offer. It has been suggested that social media sites may be the new frontier for marketing and retailing [8]. Despite the attention that Social Media have managed to attract, little is known on how they enable consumers to share product related information.

Building upon this gap, we examine how particular aspects of social media influence consumers to pass on information. We distinguish between utilitarian and hedonic motivations based on perceived value of the medium, and examine how specific characteristics of social media explain this realized value. Additionally, we perform a multi-group analysis in order to discover differentiations that may exist between males and females.

The following section presents a synopsis of literature concerning WOM and explains how social media may change the status quo of e-business through their unique features. Building upon the utilitarian and hedonic theory, in section 3 we develop a research model and formulate hypotheses regarding motivations of consumers to generate WOM. In section 4 we introduce our dataset and discuss results from the Partial Least Square (PLS) analysis. In the last section, we present conclusions derived, and highlight practical and theoretical implications of this study.

2 Theoretical Background

The importance of word of mouth has been noted in research and in practice for a long time. It is a well-known fact that positive WOM has a direct effect on increased product sales which in turn leads to further WOM. Therefore, it is critical for businesses to know what motivates users to engage in WOM, and by which means it can be facilitated. Through the digitization of WOM over the Internet, there has been a renewed interest in examining the feedback and recommendation that consumers make about products and businesses [10]. The main questions that researchers have been concentrating on are: what motivates users to WOM, and what impact positive

and negative WOM has on product sales. These two questions respectively look at the antecedents and the outcomes of WOM. Although the former still remains largely unexplored due to the contingencies to which it is subject, the later has been subject to extant research. As early as 1967, Arndt found that positive WOM impacts profitability of a firm while negative opinion sharing had the opposite effect [2]. Since then, outcomes of WOM studies have revealed that its effect extends beyond sales by increasing perceived value of a company as well as loyalty [12]. With regard to WOM motivations and antecedents, research has remained mostly a theoretical, with most research attempts grounding their hypotheses on past studies. In the current study we apply a utilitarian and hedonic motivation theory perspective in order to examine how WOM is influenced.

Utilitarian and hedonic motivation theory tries to explain the differences in consumer behaviour during the shopping process. Utilitarian and hedonic motives differ fundamentally. An action performed based on a utilitarian motivation is defined as rational and goal oriented [17]. When considering utilitarian value in the shopping context, the perceived benefit lies on the efficiency during the process of acquisition. The initiation of the shopping process for utilitarian shoppers is triggered by a mission that must be completed, and value acknowledged by the shopper is dependent on the efficiency and completeness of the process [3]. On the contrary, hedonic motivation refers to the search of emotions, such as happiness, enjoyment and fantasy, experienced during the shopping procedure. Thus, consumers propelled by hedonic motivations seek the enjoyment of the process rather than the efficiency of the process. Hence, hedonic shoppers receive satisfaction from the experience itself and the emotions which it creates.

Traditionally, researchers focused on utilitarian motives in order to explain shopping behaviours. However, recently the importance of hedonic motivation has emerged. The combined effect of these two motivations was put to test by [3], who pinpointed that hedonic factors impact unplanned shopping, while utilitarian factors do not. Therefore, it can be postulated that although the two categories differ significantly, in order to view the shopping processes holistically, both must be investigated simultaneously. Antecedents that form hedonic and utilitarian motivation, as well as differences in shopping behaviour have also been subject of much attention over the past years [25]. These outcomes provide interesting implications since the behaviour of utilitarian and hedonic shoppers differ in many ways.

A number of recent scientific articles have examined these types of motivations with respect to the design of online environments [22]. Most such studies have identified utilitarian factors as being the primary force for shopping online [1]. Despite the dominance of utilitarian factors in influencing shopping behaviour, research on the hedonic aspects has not remained stagnant. [20] argue that consumers do not navigate through electronic shops solely to collect information about a specific product, but also to fulfill the needs of experience and emotion. Similarly to traditional means of shopping, the actions performed during the shopping process are influenced by the aesthetics and enjoyment that it provides [23]. Summarizing the above, it can be asserted that with regard to shopping, utilitarian and hedonic motivations exist regardless of the context in which they are promoted.

We can identify a broadening in perspective with the terms hedonic and utilitarian not only being applied to motivations, but also to perceived usefulness of systems and aspects of experience [26]. A number of studies have gone past the examination of motivations as predictors of product search intention and purchasing intention, and applied the concepts in the domain of Information Systems (IS) [32], software design [15] and mobile device usage [33]. However, little research has been performed so far on the motivations for browsing products on social media platforms through the utilitarian and hedonic motivation lens [25]. Additionally, most studies focus on antecedents of utilitarian and hedonic motivation, and with regard to shopping outcomes, only examine intentions to purchase. We argue that identifying how perceived utilitarian and hedonic value impacts WOM is an equally important question based on the growing literature. In this study, we investigate how perceived utilitarian and hedonic value affects WOM and explain this effect with the use of two mediated constructs: customized advertisements and socializing ability.

3 Research Model

3.1 Utilitarian Motivation

Utilitarian motivations are regarded in literature as one of the two types of forces that engage consumers in the buying process [31]. In fact, it has been suggested that users with a strong utilitarian motivation, will both browse and purchase items that they are looking for [31]. Going beyond examining the purchasing intentions, [18] showed that users propelled by utilitarian motives will also end up engaging in word-of-mouth with peers.

A utilitarian motivation is not solely a result of the product itself, but also of the medium that is used to promote it [22]. This places emphasis on the features that should be considered when designing it. Taking into account that the first step in the purchasing process is the seeking of information about a product, whether goal-oriented or exploration-oriented, and in conjunction with the importance that is noted regarding the platform which it is promoted, we hypothesize the following regarding social media:

H1. *Utilitarian motivation of social media users (for product browsing) has a positive impact on Word-of-Mouth.*

One of the strongest points of social media in relation to marketing is their ability to generate advertisements that are tailored to the likings of each consumer. Social media users continuously declare their likings by joining product groups, following company news, submitting personal information on their online profiles etc. This constitutes social media as a very effective medium for presenting products and services to consumers and riding them of the information overload hassle. In many occasions this is not done in a direct way, i.e. presenting an advertisement in a panel of a screen, but can be also accomplished in a more pervasive manner. According to [30], the service of customized advertisements is a feature that adds utilitarian value to a

medium, meaning that the process of searching becomes more efficient. Research has identified that the relevance of customized advertisements is a strong predictor for the intention to purchase a product/service [7]. Additionally, it has been shown that features such as customized advertisements, add utilitarian value to the shopping process and have an impact on word of mouth [18]. Especially in the case in where there is a high degree of homophile, i.e. similarity of likings between users or information that is presented to them, user`s tend to proceed to WOM. Therefore, we hypothesize that:

H2. *Customized advertisements on social media mediate the positive influence that utilitarian motivation has on word-of-mouth.*

3.2 Hedonic Motivation

Although the impact of hedonic motivations as part of the shopping process was noted late in comparison with utilitarian motivations, they are now considered as an essential aspect to be examined. Hedonic features have a greater impact on exploration-oriented consumers, which if sufficiently stimulating result in unplanned purchasing of products [31]. Accordingly, hedonic aspects are of equal importance when considering mediums of product browsing and commerce. Although marketers typically provide web-stores that fulfill all functional requirements, it is emphasized that they should not neglect more escapist items [17]. The authors note that such elements that are intrinsically enjoyable facilitate greater user participation. This engagement perceived by users is also found to impact user's word-of-mouth intention for conventional stores [18]. Hence, we can assume that for social media the same will apply:

H3. *Hedonic motivation of social media users (for product browsing) has a positive impact on Word-of-Mouth.*

Hedonic motivated shoppers are inclined to engage in the process of browsing and purchasing for a number of reasons. However, one of the most frequently noted motivations is the ability to engage in social interaction with others. This is an aspect noted as being a major hedonic motivation for product browsing and purchasing in conventional stores [28]. Through the emergence of Web 2.0 applications and especially social media, users can transfer this experience in an online community in a synchronous or/and asynchronous manner. According to [34], engaging in online conversations during the browsing process could elevate their perceived hedonic value, and motivate them to purchase and browse in the future. When individuals with common likings interact online and share opinions, it is found that engaging in WOM is enhanced. Therefore, social interaction in online communities or chat has an impact on purchasing, which in turn leads to WOM. This WOM that is facilitated then results in additional purchases, thus making these two concepts interlinked [4]. Consequently we can assume that:

H4. *The socializing ability of social media mediates the positive influence that hedonic motivation has on word-of-mouth.*

Table 1. Construct Operational Definitions

Construct	Operational Definition	Sources
Customized Advertisements	*The perceived value from the service of customized advertisements presented to users of social media based on their likings*	[14], [31]
Socializing Ability	*The ability to engage in a conversation (synchronous or asynchronous) with fellow peers during the browsing of products on social media sites*	[31], [34]
Utilitarian Motivation	*The degree to which users perceive social media to be a useful and effective means to browse products.*	[6], [26]
Hedonic Motivation	*The degree to which users perceive browsing products on social media to be a fun and emotionally stimulating experience.*	[6], [26]
Word of Mouth Intention	*The intention of social media users to share information with peers about products or services which they find relevant or interesting through such platforms.*	[4], [16]

4 Data Collection and Measurements

For each of the constructs described in the conceptual model a number of items were adapted from past and were measured by applying a 7-point likert scale. The final questionnaire was divided into three parts; the first gathered information about the demographics of the sample, the second measured the utilitarian and hedonic motivation constructs, while the third targeted in measuring intention to share information. Questionnaire links were e-mailed to social media users through a mailing list of over 600 respondents of an institution, and posted on 20 forum boards of various themes (Fashion, computers, hobbies etc.). The questionnaire was active for a period of 7 months approximately, from December 2011 to June 2012. In total 169 responses were received and retained for further analysis with a valid response rate of 16,9% (from the mailing list).

Table 2 presents some descriptive results regarding our sample. Only respondents that maintained at least one social media account were qualified to participate in the research. An almost equal amount of male and female respondents filled out the questionnaire, with younger user groups being more active in social media sites. The distribution of ages is justified since the vast majority of social media users also belong to these age categories (15-44). With respect to online spending habits, the vast majority is accustomed to purchasing; therefore they exhibit familiarity with the online buying process. Finally, outcomes from respondent's replies confirm recent suggestions that social media users are considerably active by revealing that 76% of them visit them on a daily basis.

Table 2. Sample Descriptive Statistics

Variable		*%(N=169)*	*Variable*		*%(N=169)*
Gender	Male	56%	**Online Expenditure (€/Month)**	Never Purchased Online	6%
	Female	44%		1-24 €	44%
Education	Primary School	1%		25-50 €	25%
	Junior High School	1%		51-100 €	16%
	High School	14%		101-250 €	7%
	University	56%		> 250 €	2%
	Post Graduate	28%	**Frequency of SM use**	Several times a day	55%
Age	<24	50%		About once a day	21%
Group	25-34	37%		3-5 times a week	8%
	35-44	7%		1-2 times a week	4%
	45-54	6%		Every few weeks	4%
	>55	0%		Less often	7%

In accordance with the grouping of social media sites as proposed by [19], we developed a list of websites from which users could those that they maintained an active presence in. Furthermore, we included the option for additions of social media websites which we had not included initially by users and after examining if indeed they belonged to the social media spectrum where either kept or omitted from the list. As results indicate, the most popular websites among our group of respondents are YouTube (92%), followed by Facebook (82%) and Wikipedia (74%).

5 Analysis

Partial least squares (PLS) path modelling is applied in order to validate the causal relationships amongst variables as defined in the hypotheses. The minimum amount of observations according to Smart PLS documentation is ten for each of the constructs, thus the sample of 169 users surpasses this threshold.

The validity and reliability of the items was examined by testing item loadings are above the threshold of 0.7. Since loadings were well above 0.7, we can be sure that we have reliable indicators. Additionally, convergent validity was confirmed according to the proposed methodology for PLS by [13] by checking that Average Variance Extracted (AVE) values were all greater than 0.5.

In sequence, constructs were examined to determine if they are reliable to work upon. Composite reliability values measuring the internal consistency of each construct, were all above 0.7, and are therefore considered reliable [13]. Additionally, the

Average Variance Extracted (AVE) values were all greater than the squared correlations for any other construct, thus confirming discriminant validity. The outcomes of the discriminant and convergent validation tests are depicted in Table 3, with AVE being depicted in bold on the diagonal of the matrix.

Table 3. Inter-Construct Correlations and Validity Measures

Items	Mean	S.D.	CR		UT	HE	CA	SC	WOM
5	3.915	1.750	0.939	**UT**	**0.739**				
5	3.348	1.896	0.959	**HE**	0.722	**0.826**			
3	3.096	1.907	0.908	**CA**	0.284	0.278	**0.768**		
3	4.051	1.963	0.890	**SC**	0.490	0.485	0.494	**0.730**	
5	3.071	2.017	0.922	**WOM**	0.453	0.463	0.503	0.565	**0.702**

The hypothesized model represents a mediation model. In order to test if the mediator variable fully mediates the association between independent and outcome variable, two effects are put to test. First we examine the direct effect of utilitarian and hedonic motivation on WOM. Then we test the relationship between independent variables is mediated. In order to test these effects, a bootstrapping procedure with replacement was used taking 1000 subsamples to estimate statistical significance. Figure 1 summarizes the results of the structural model assessment of the mediated analysis.

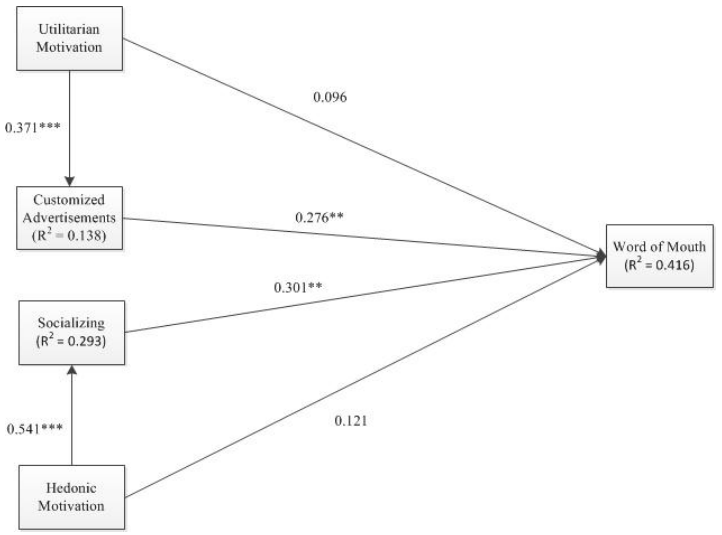


Fig. 1. Structural Model with Path Coefficients

To test the first hypothesis, we examine the direct effect that utilitarian motivation has on WOM without the inclusion of a mediator variable ($\beta = 0.313$, $p < 0.01$). This indicates that there is a positive and significant effect of utilitarian motivation on WOM, thus confirming hypothesis 1. To test hypothesis 2 we follow the mediation testing procedure proposed by [11], including the mediator construct. Utilitarian motivation has a positive and significant effect on customized advertisements ($\beta = 0.371$, $p < 0.001$), and customized advertisements has a positive and significant impact on WOM ($\beta = 0.276$, $p < 0.01$). To test if the mediation is significant, we perform Sobel`s mediation test which has a significant value of 2.345 ($p < 0.01$), hence confirming hypothesis 2 [29]. Since the effect of utilitarian motivation on WOM in the mediated model is found to be non-significant ($\beta = 0.096$ $p > 0.05$), we can state that customized advertisements fully mediate this relationship.

For hypothesis 3, the result of the direct effect of hedonic motivation on WOM yields a positive and significant effect ($\beta = 0.260$, $p < 0.01$). Therefore, we accept the statement that hedonic motivation influences WOM in a positive manner. In the mediated model, hedonic motivations effect on the mediator (Socializing) is significant and positive ($\beta = 0.541$, $p < 0.001$), as is the impact of the mediator to WOM ($\beta = 0.301$, $p < 0.01$). The total effect of the mediation is significant according to the Sobel test statistic with a value of 2.593 ($p < 0.01$), confirming hypothesis 4. In total, the model explains 41.6% of variance for the WOM construct, which is at moderate levels. The direct effect of hedonic motivation to WOM is rendered as non-significant in the mediated model ($\beta = 0.121$, $p > 0.05$), which means that socializing ability fully mediates the effect.

When performing a multi-group analysis splitting the sample into male and female we notice that a deviation from our results occurs for male users. The effect that perceived value of customized advertisements has on WOM is found to be positive but non-significant ($\beta = 0.209$, $p > 0.05$), while the effect of the mediation in total has reduced significance according to the Sobel test statistic (1.746, $p < 0.05$).

Table 4. Effects of associations and significance levels

	Entire Sample	**Male**	**Female**
Model (Direct Effect)			
UT -> WOM	0.313**	0.297**	0.341**
HE -> WOM	0.260**	0.372**	0.322**
R^2	**0.284**	**0.369**	**0.366**
Model (Mediated)			
UT -> CA	0.371***	0.399***	0.371***
CA -> WOM	0.276**	0.209	0.290**
UT-> CA -> WOM (Sobel Statistic)	2.345**	1.746*	2.475**
UT -> WOM	0.096	0.052	0.220
HE -> SC	0.541***	0.648***	0.497***
SC -> WOM	0.301**	0.417**	0.262**
HE -> SC -> WOM (Sobel Statistic)	2.593**	3.449***	2.646**
HE -> WOM	0.121	0.119	0.021
R^2	**0.416**	**0.463**	**0.401**

6 Discussion and Implications

Our research aim was on explaining with the aid of utilitarian and hedonic theories, what features of social media affects users intention to WOM. In particular, we examine how these motives are facilitated on social media sites, and what features trigger this effect. To this end, we hypothesize that the perceived utilitarian and hedonic value of social media sites enhances WOM through the features of customized advertisements and socializing ability. Empirical outcomes support these two theoretical contributions.

The first contribution is an explanation of the importance of utilitarian motivation. In the initial model we show that there is a strong relationship between the utilitarian motivation of users on social media and their intention to pass on information. Through the mediating construct of customized advertisements we attempt to show that the utilitarian motive is facilitated through this feature. Traditionally, the effectiveness of customized advertisements is measured by monitoring the change in sales of a product. However, recent studies have noted the significance that WOM has, since it may result in an incremental increase of sales. Our results demonstrate that social media websites, due to their ability to present products or services that are in accordance with user's likings, will positively affect sharing of information. Since customized advertisements fully mediate the relationship of utilitarian value with WOM, we can conclude that they are the prime feature in motivating users to pass on product related information.

The second contribution concerns how the hedonic value perceived when browsing products on social media leads to WOM. We use the construct of socializing ability to measure the tendency of users to engage in social interaction while browsing. We find that when users interact with peers that are interested in the same products/companies, they are more inclined to share that information with their friends.

Consequently, we conclude that utilitarian and hedonic motivations result in WOM through the value perceived from the customized advertisements, and the ability to engage in social interaction. These two features, although highly significant in explaining WOM, may not be the only ones, since outcomes reveal that 58% of variance still remains unexplored. These could be other features on social media sites of a utilitarian or hedonic nature, or even be related to product aspects.

Another interesting observation concerns the differences between genders. From our split-sample analysis, we identified that male users are not motivated to engage in WOM because of the perceived value of customized advertisements. Instead, they are more inclined to do so compared with females due to the perceived value of socializing. In the initial model however, no significant differentiation is noted between utilitarian and hedonic motivation between genders. This finding raises implications for practitioners, whom should proceed to different strategies based on gender when targeting to participate in WOM. Therefore, managers should be well aware of these contingencies when constructing and deploying a marketing strategy over social media.

Our study presents a number of limitations that future research may address. One of the most important limitations is that we examined social media as a whole,

including a wide array of sites. In order to have a holistic view of how WOM is spread over social media, they may need to be examined separately. We chose to focus on the broad term of social media, since this study is one of the first to explore motivations in such a manner. We encourage future researchers to examine additional factors, specific to social media, which may influence WOM and complement our model.

References

1. Alba, J., Lynch, J., Weitz, B., Janiszewski, C., Lutz, R., Sawyer, A., Wood, S.: Interactive home shopping: consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. The Journal of Marketing, 38–53 (1997)
2. Arndt, J.: Role of product-related conversations in the diffusion of a new product. Journal of marketing Research, 291–295 (1967)
3. Babin, B.J., Darden, W.R., Griffin, M.: Work and/or fun: measuring hedonic and utilitarian shopping value. Journal of Consumer Research, 644–656 (1994)
4. Brown, J., Broderick, A.J., Lee, N.: Word of mouth communication within online communities: Conceptualizing the online social network. Journal of Interactive Marketing 21(3), 2–20 (2007)
5. Chen, Y., Xie, J.: Online consumer review: Word-of-mouth as a new element of marketing communication mix. Management Science 54(3), 477–491 (2008)
6. Childers, T.L., Carr, C.L., Peck, J., Carson, S.: Hedonic and utilitarian motivations for online retail shopping behavior. Journal of Retailing 77(4), 511–535 (2002)
7. Choi, S.M., Rifon, N.J.: Antecedents and consequences of Web advertising credibility: A study of consumer response to banner ads. Journal of Interactive Advertising 3(1), 12–24 (2002)
8. Constantinides, E., Romero, C.L., Boria, M.A.G.: Social media: a new frontier for retailers? European Retail Research 22(1), 1–28 (2008)
9. Darden, W.R., Reynolds, F.D.: Shopping Orientations and Product Usage. Journal of Marketing Research 8, 505–508 (1971)
10. Dellarocas, C.: The digitization of word of mouth: Promise and challenges of online feedback mechanisms. Management Science 49(10), 1407–1424 (2003)
11. Edwards, J.R., Lambert, L.S.: Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. Psychological Methods 12(1), 1 (2007)
12. Gruen, T.W., Osmonbekov, T., Czaplewski, A.J.: eWOM: The impact of customer-to-customer online know-how exchange on customer value and loyalty. Journal of Business Research 59(4), 449–456 (2006)
13. Hair, J.F., Ringle, C.M., Sarstedt, M.: PLS-SEM: Indeed a silver bullet. The Journal of Marketing Theory and Practice 19(2), 139–152 (2011)
14. Han, J., Han, D.: A framework for analyzing customer value of internet business. Journal of Information Technology Theory and Application (JITTA) 3(5), 4 (2001)
15. Hassenzahl, M., Platz, A., Burmester, M., Lehner, K.: Hedonic and ergonomic quality aspects determine a software's appeal. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2000)
16. Hennig-Thurau, T., Gwinner, K.P., Walsh, G., Gremler, D.D.: Electronic word‐of‐mouth via consumer‐opinion platforms: What motivates consumers to articulate themselves on the Internet? Journal of Interactive Marketing 18(1), 38–52 (2004)

17. Hirschman, E.C., Holbrook, M.B.: Hedonic consumption: emerging concepts, methods and propositions. The Journal of Marketing, 92–101 (1982)
18. Jones, M.A., Reynolds, K.E., Arnold, M.J.: Hedonic and utilitarian shopping value: Investigating differential effects on retail outcomes. Journal of Business Research 59(9), 974–981 (2006)
19. Kaplan, A.M., Haenlein, M.: Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 53(1), 59–68 (2010)
20. Kim, Y.M., Shim, K.Y.: The influence of Internet shopping mall characteristics and user traits on purchase intent. Irish Marketing Review 15, 25–34 (2002)
21. Knapp, M.L., Daly, J.A.: Handbook of interpersonal communication. Sage Publications, Incorporated (2002)
22. Kourouthanasis, P.E., Giaglis, G.M., Vrechopoulos, A.P.: Enhancing user experience through pervasive information systems: The case of pervasive retailing. International Journal of Information Management 27(5), 319–335 (2008)
23. Mathwick, C., Malhotra, N., Rigdon, E.: Experiential value: conceptualization, measurement and application in the catalog and Internet shopping environment. Journal of Retailing 77(1), 39–56 (2001)
24. Mikalef, P., Giannakos, M., Pateli, A.: Exploring the Business Potential of Social Media: An Utilitarian and Hedonic Motivation Approach. In: Proceedings of the 25th Bled e-Conference, Bled, Slovenia (2012)
25. Mikalef, P., Giannakos, M., Pateli, A.: Shopping and Word-of-Mouth Intentions on Social Media. Journal of Theoretical and Applied Electronic Commerce Research 8(2) (2013)
26. O'Brien, H.L.: The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. Interacting with Computers 22(5), 344–352 (2010)
27. Richins, M.L., Root-Shaffer, T.: The role of evolvement and opinion leadership in consumer word-of-mouth: An implicit model made explicit. Advances in Consumer Research 15(1), 32–36 (1988)
28. Rook, D.W.: The buying impulse. Journal of Consumer Research, 189–199 (1987)
29. Sobel, M.E.: Asymptotic confidence intervals for indirect effects in structural equation models. Sociological Methodology 13, 290–312 (1982)
30. Sun, Y., Lim, K.H., Jiang, C., Peng, J.Z., Chen, X.: Do males and females think in the same way? An empirical investigation on the gender differences in Web advertising evaluation. Computers in Human Behavior 26(6), 1614–1624 (2010)
31. To, P.L., Liao, C., Lin, T.H.: Shopping motivations on Internet: A study based on utilitarian and hedonic value. Technovation 27(12), 774–787 (2007)
32. Van der Heijden, H., Verhagen, T., Creemers, M.: Understanding online purchase intentions: contributions from technology and trust perspectives. European Journal of Information Systems 12(1), 41–48 (2003)
33. Wakefield, R.L., Whitten, D.: Mobile computing: a user study on hedonic/utilitarian mobile device usage. European Journal of Information Systems 15(3), 292–300 (2006)
34. Wolfinbarger, M., Gilly, M.C.: Shopping online for freedom, control, and fun. California Management Review 43(2), 34–55 (2001)

Analysis of Social Network Dynamics with Models from the Theory of Complex Adaptive Systems

Ilias Lymperopoulos and George Lekakos

Department of Management Science and Technology
Athens University of Economics and Business, Athens, Greece
{eliasliber,glekakos}@aueb.gr

Abstract. The understanding and modeling of social dynamics in a complex and unpredictable world, emerges as a research target of particular importance. Success in this direction can yield valuable knowledge as to how social phenomena form and evolve in varying socioeconomic contexts comprising economic crises, societal disasters, cultural differences and security threats among others. The study of social dynamics occurring in the aforementioned contexts with the methodological tools originating from the complexity theory, is the research approach we propose in this paper. Furthermore, considering the fact that online social media serve as platforms of individual expression and public dialogue, we anticipate that their study as complex adaptive systems, will significantly contribute to understanding, predicting and monitoring social phenomena taking place on both online and offline social networks.

Keywords: Social dynamics, social network analysis, complex adaptive systems, online social networks, online social media, offline social networks, complex networks, complexity theory, scale free networks, statistical physics, complex adaptive systems models.

1 Introduction

In a society where social interactions obtain a global dimension due to the increased mobility of people and proliferation of social networking tools which keep citizens in a perpetual contact, complexity and unpredictability rises at a rate faster than ever before. In this context, the analysis and modeling of social dynamics is being redefined as new approaches are introduced in studying the social systems which become all the more less-bounded, nonlocal and complex. In such a setting, nonlinear phenomena emerge on social networks as the parameters regulating the individual's behavior can bring the system to its "tipping points" where abrupt phase transitions take place thus indicating the sensitivity of social behavior to small changes in the parameters that control the system.

As the social networks become more complex, the analytical tendencies shift from the study of static structural and locational properties to the analysis of the temporal change in their structure and composition and its impact to ties among individuals and social norms. Furthermore, new thinking concepts are applied when analyzing their

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 124–140, 2013.

structure as it is considered the result of a self-organization process giving rise to highly connected social clusters playing a crucial role in terms of social stability and transmission of interaction. Also, the interplay between individuals' state and social network topology causes a host of complex social dynamics such as opinion formation, spread of ideas, influence, epidemics and community formation among others.

This new theoretical framework originating from complexity theory sets the study of social dynamics on a new basis. As a result, we propose that social behavior phenomena stemming from economic crises, societal disasters, cultural differences and security threats, be studied and modeled as "emergent properties" deriving from simple forms of local social interactions spawning global effects. In that vein, trust, confidence, influence and persuasion can also be cast in complexity theory terms and modeled as outcomes of the interplay between nodes' state and social network topology. To this end, the analysis of information obtained from online social media which serve as platforms of individual expression and public dialogue, will significantly contribute to understanding, predicting and monitoring social dynamics occurring on both online and offline social networks.

The paper is organized into six sections. After the brief introduction, section two provides a review of methodologies applied to structural and locational analysis of social networks outlining the type of problems these tools are suitable for, meanwhile identifying their weaknesses in dealing with the analysis of the temporal dynamics taking place on them. Section three introduces network models appropriate for studying the co-evolutionary adaptation of network dynamics and topology in social networks. Section four focuses on significant topological properties common among a large number of complex networks and illustrates why offline and online social networks can be analyzed as complex adaptive systems. Section five presents models originating from the complex adaptive systems theory used for studying social phenomena, thus providing evidential proof of suitability for analyzing social dynamics occurring on online social networks. The paper closes with the concluding remarks in the sixth section.

2 Structural and Locational Analysis of Social Networks

Tracing the beginnings of the Social Network Analysis scientific field, we go back in the early 1930's when Moreno [1] introduced the sociogram - a new methodological technique which was described as "invention" since it was the first systematic attempt to plot the structure of social relations in a group. This innovation heralded the advent of sociometry and social psychology which established qualitative methods for exploring socio-emotional networks [2]. Since then, much progress has been achieved with the theoretical notions of communities, popularity, prestige, transitivity, clique, social role, reciprocity, influence, dominance and conformity, providing impulse for the development of social network analysis methods [3].

Tools from network theory have been used for the description and analysis of social networks. This methodological approach was initiated by anthropologists such as Mitchell and Barnes who noticed that the analysis and understanding of the

people's behavior in societies of increasing complexity, as the population was moving from rural to urban areas, could not be performed under the traditional concepts of social institutions such as economics, religion, politics e.t.c. The application of network theory to social network analysis, was further expanded by Wasserman's and his colleagues research centered around the principle that social phenomena share an underlying relational network structure and therefore they can be analyzed with the theoretical framework of graph theory [2, 3].

The application of graph theory in the social network analysis field provided a powerful set of tools for describing and modeling the relational context in which behavior takes place, as well as the relational dimensions of that behavior. Graph theory was also applied in the investigation of associations among concepts and developmental phenomena such as the structure of personal life histories. It was also the basis for the development of methods addressing the identification of cohesive groups, blockmodeling and equivalence analysis, dynamical network analysis, structural balance and methods for the analysis of two mode data (e.g. person by event). Approaches like the foregoing ones rest on the capability of graph theory to systematically measure the properties of a social structure in a way that renders modeling and comparison smoother [4]. With their capacity to analyze network structural properties, the methodological tools supplied by graph theory have been successfully applied in predicting work satisfaction and team performance [5], power and influence [6], success in bargaining in a competitive context [7, 8], mental health outcomes and a variety of other social phenomena. In the recent years the methodologies based on graph theory have evolved as powerful instruments for structural and locational analysis of social networks [2, 3, 9–12].

However, despite the successful application of graph theory in social network analysis tasks pertaining to their descriptive features, the analytical focus is limited to static network properties. This downside has begun to being addressed lately as the emphasis shifts from analyzing static properties to analyzing dynamic processes taking place on social networks [13].

In tandem with the shift of focus on the dynamic properties of social networks structure, efforts have been made to employ basic statistical measures of probability and significance to assess the validity of propositions about network structure [14]. However despite the analytical power of standard statistical tools such as regression, significance tests and variance analysis, their application to network data comes with drawbacks originating from the hypothesis that network observation data is independent, an assumption that contradicts with typical network data. This inefficiency has been addressed by Stanley Wasserman and his colleagues [15, 16] with the introduction of the Exponential Random Graph Models (often referred as p* models). These models randomly produce graphs on a given set of points covering the entire range of possible networks from completely unconnected to fully connected. The log odds ratios of the probability of each randomly generated graph are used for producing Monte Carlo estimations that allow the comparison of a real network with the set of networks previously created by the model so that it can be estimated how possible is that the actual network can exist only by chance. This modeling technique can be used for validating theories about network structures and relational process. In other words the implemented model puts forward specific structural results which are then

compared with the real network data taken from observations. To exemplify the concept of this modeling technique we can hypothesize that mutuality and cooperation tendencies appear frequently among groups with common interests in a community. For assessing the validity of this proposition we can construct an ERGM statistical model incorporating the hypothesis and subsequently test whether statistical results are confirmed by empirical data.

In general, the main fields of social network research concern: (i) the study of the static topological properties, (ii) the modeling of network formation (either static or growing), and (iii) how the dynamical processes occurring on networks are affected by the network topology [17].

Nevertheless, when trying to reproduce the topological properties of real networks by means of simple graph models there exist difficulties in deriving certain properties often common between networks of different nature thus indicating the existence of common organizing mechanisms [17]. The limitations –inherent in simple graph models- have been the subject of recent research focusing on the temporal aspects of networks growth as well as on the dynamics occurring on local and global level. The next section of this paper serves as a synopsis of the developments on this field meanwhile proposing new types of network models for the analysis of both offline and online social networks.

3 Temporal Dynamics of Social Networks

The foregoing approaches serve the structural and locational analysis of social network properties. However, for investigating the temporal network dynamics new concepts such as the small world [18, 19], preferential attachment [20, 21], percolation [22, 23] and epidemiology [24–26] are used as they better represent the structure and growth dynamics of social networks showing the following nontrivial topological properties: (i) Correlation (or anticorrelation) between degrees of neighboring nodes leading either to establishment of a link (assortativity) or the avoidance of the connection (dissartotativity) [12, 26, 27], (ii) Small World effect [18, 19] and (iii) High Clustering Coefficient [20].

Since the social network evolution is based on the co-evolutionary adaptation of dynamics and topology by local rewiring rules, a question emanating from this observation is whether co-evolution of dynamics extends to a global time scale despite the local nature of the rewiring events and a significant time scale separation between dynamical and topological updates [28]. Focusing on the interactions between the nodes of a social network rather than the attributes of the nodes themselves for deriving insights as to the dynamics taking place on social networks, an interdisciplinary perspective is being formed encompassing the scientific fields of statistical physics, nonlinear dynamics, critical phenomena, fractal geometry, spin glasses, and many-body theory [17].

It is asserted that the complexity of real-world networks stems from the interplay between the network topology and network dynamics (figure 1). Several studies up until now concentrate either on the effect that topological properties have on network

dynamical processes or the opposite, that is the effect of node-specific dynamical variables on network structure [17]. As a result an interesting line of research is the investigation of the combined effect of these processes on each other. Such an approach entails that we should examine the network dynamical processes and the network growth as happening in the same timescale and not on separated ones in which case the slower variables enslave the faster ones. The simultaneous execution of these processes give rise to a new class of networks which are described with the term Complex Adaptive Systems.

According to the definition provided by G.M. Murray, "A complex adaptive system acquires information about its environment and its own interaction with that environment, identifying regularities in that information into a kind of 'schema' and acting in the real world on the basis of that schema" [29]. This means that a Complex Adaptive System (CAS) derives from its surrounds which at the same time is affected by the CAS itself. The complexity of these systems is attributed to the dynamic nature of interactions among the network nodes giving rise to system properties which cannot be handled as aggregations of the properties of the individual static entities. These systems are adaptive because the individual and collective behavior mutates and self-organizes in response to a triggering micro-event or series of events.

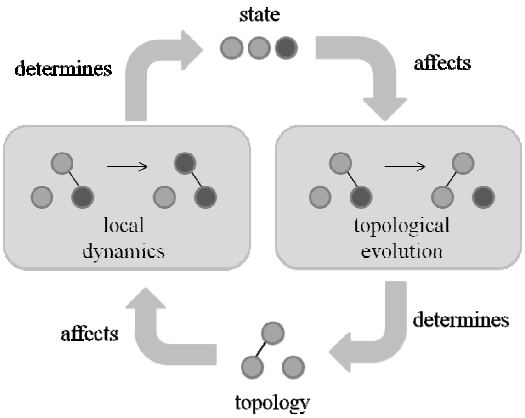


Fig. 1. In an adaptive network the global topology and the dynamics of the nodes are coupled with a feedback loop[30]

The complexity of a social network results from the dynamic feedback loops inherent in the human behavior in a societal context. These feedback loops cause the system to behave nonlinearly. It is really astonishing how fast a social network can pass to the complexity realm. The connections among individuals who predict and react to the predictions and actions of the others coupled with the progressive exacerbation of these actions as the system becomes all the more connected, leads to nonlinear interactions difficult to decompose and then complexity sets in [31]. For instance, in the opinion formation process, the beliefs of each individual node of the social network, and the evolution of the topology of the network are blended in a way that people's

connectivity affects their opinions and at the same time their opinions affect their connections [32].

The adaptive dynamics is introduced in a social network due to the ability of the individuals to control the number, type and duration of their interactions with others [33]. After a large number of adaptive cycles, evolution towards a self-organized critical state is observed. An adaptive coevolutionary scheme leading to network self-organization is based on the theory that correlated activity connects, whereas decorrelated activity disconnects [28]. This coevolutionary pattern is also found in the neural network theory which posits that if the activity of two neighboring neurons is on average highly correlated then they will obtain a common link. However if their activity is on average less correlated, then they will lose their common link [28]. The Hebbian learning shows that network self-organization by correlation-driven rewiring is robust even when spatial constraints are present and dynamics is affected by noise [34]. The phenomenon of self organized criticality constitutes an important feature of adaptive networks and boils down to the fact that each node in a network is recipient of dynamical information which relates to the connectivity of the entire network, thereby providing global information to individual nodes thus causing the network to organize itself [17, 28].

Social networks are self-organized systems in that they consist of many interacting entities co-operating for achieving the desired result. The structure of a self-organizing system emerges without explicit external force and is the result of internally enforced constraints generated by the interactions among the entities. A self-organized system evolves temporally and spatially thus exhibiting varying organizational structures. The actions taking place in a self-organized system result from the negative and positive feedback mechanisms by means of which the system gears towards stability or chaos respectively [31].

Adaptive social networks result from the combination of contact processes with rules for network topological evolutions. Contact processes refer to how individuals are changed and shaped through interactions with others. For example, present models study the information spreading as a result of two rival process that is social adjustment (contact process) and social segregation (topological evolution) [35]. Holme and Newman report that the diversity of opinions coexisting in a society go through a phase transition if the relative rate of social adjustment and social segregation crosses a critical value [36]. When social segregation is coupled with social adjustment this leads to an adaptive network [35].

As complex adaptive systems social networks have to cope with the contradictory needs of accommodating response to changing environmental signals, while maintaining a sufficient level of stability in the dynamical networks that process this information. This remark gives rise to the idea that complex adaptive systems may have evolved to the “edge of chaos” between ordered and disordered dynamical regimes [37].

The consideration of adaptive social networks is an important step towards more realistic models of social interactions in structured populations. Coupling the dynamics on networks with the dynamics of networks leads to emergent new phenomena outside the classical context of social dynamics on static networks [33]. However a

complete understanding of the behavior of adaptive networks requires tools which encompass stochastic dynamics, topology and time dependent graph theory [38].

4 The Structure of Social Networks

Over the past years researchers have found that a large number of complex systems share some important common properties which hold a prominent role in the way complex systems are studied [20]. In brief these properties refer to:

Small worlds: This concept outlines the fact that despite the large size of complex networks, the path between any two nodes is relatively short [18].

Clustering: It is often observed in complex networks that there are fully connected sub-graphs. In social networks this is called "cliquishness" and refers to circles of friends where each individual knows each other. The propensity of social networks to cluster is quantified by the clustering co-efficient metric [18, 19].

Scale free structure: The degree distribution shows a scale free structure that is a heterogeneous network topology encompassing a relatively small number of nodes with exceptionally high degree and a very large number of nodes with low degree (figure 2). One remarkable feature of scale-free networks is that they are highly robust against random errors but on the contrary they are highly vulnerable to attacks targeting nodes which serve as hubs in the network. The scale free modeling lays the emphasis on the network dynamics and its main objective is to model the process that underpins the evolution of complex networks since such an approach will lead to the creation of networks with the correct structure [20, 25].

Fig. 2. Scale Free network: Small number of nodes with high degree and a large number of nodes with low degree

The aforementioned features have been observed in a remarkable number of networks including the internet, the World Wide Web, scientific collaboration, actors, citation, business, molecular, biological and social networks among others [39, 40]. Recently, empirical results show that online social networks encompass the foregoing

properties. Mislove et al. [41] studied the online social networks Flickr, LiveJournal, Orkut and Youtube and their results show that these networks show the small world, high clustering and scale free properties observed in offline social networks. Catanese et al. [42] studied the Facebook friendship relations, and discovered that Facebook also comprises the small world, high clustering and scale free properties featured in offline social networks. These studies confirm that online social networks contain a large, densely connected cluster of high-degree nodes, bordered by lots of small clusters of low-degree nodes. This implies that high-degree nodes are crucial for the connectivity network and the flow of information which in social systems is content dependent, meaning that different type of information may be spreading over different social networks adhering to different dynamical rules [43].

From the above and similar considerations, becomes evident that an in-depth understanding of the social dynamics occurring on online social networks can yield significant insights as to how similar social phenomena take their course on off-line social networks. As a result, the analysis of online social networks from the perspective of Complex Adaptive Systems emerges as an innovative approach by means of which social phenomena can be better understood and modeled with wide scale socioeconomic benefits as predictability and monitoring power is gained regarding social transformations brought about by economic crises, disasters, security threats, criminality, epidemics and globalization to name but a few pressing issues. In this theoretical framework the study of the intermediate-scale substructures in networks, consisting of vertices more densely linked to each other than the rest of the network, has become an outstanding research subject in the complex network theory [44].

Models originating from the Complex Adaptive Systems theory which can be used for effecting the foregoing research objectives are presented in the next section. These models are able to capture the dynamic nature of social phenomena thus providing distinctive analytical advantages when it comes to dealing with the temporal aspects of interactions in social networks consisting of heterogeneous actors.

5 Studying Social Phenomena with Complex Adaptive Systems Models

The traditional models of studying social networks concentrate on static, homogeneous situations comprising few or a large number of agents in an environment where time and space are not of importance. However, studying the social networks with static models falls short of capturing the dynamic nature of interactions. Also, a key driving force in social networks is heterogeneity which gives rise to a rich set of alternative behaviors. With that said, traditional social networks analysis focusing on average behaviors, not only might be incomplete, but also illusory [31]. In order to overcome these shortcomings, a new category of network models has been developed with a view to capturing in a more sufficient way the complexity of many physical systems including the social ones (table 1). In the complex adaptive systems models, the topology changes dynamically in reaction to changing node characteristics and as a result the topological changes trigger changes in the dynamics occurring on nodes.

The main characteristic of complex adaptive networks is the interplay between node dynamics and network topology meaning that nodes and links evolve over time even when a steady state is reached [45]. The scientific field of complex adaptive systems by being able to explore the dynamical aspects of the systems, is relevant for analyzing some of the most urgent social issues such as economic crises, globalization, sustainability, fighting terrorism and preventing epidemics [31].

Table 1. The new Social Network Analysis concepts focus on the dynamic nature of interactions and the heterogeneity of agents

Social Network Analysis	
Traditional Models	**CAS Models**
Study of Static Situations	Study of Dynamic Situations
Homogeneous agents	Heterogeneous Agents

Studying social networks under the lens of complex networks theory entails the application of modeling techniques complying with the concept that the network topology develops by means of a feedback mechanism which connects topology and dynamical processes running on the network itself. Deciphering the network structure is actually an indispensable step in investigating dynamical processes occurring on social networks. For example, It has been shown that the network structure plays a vital role in the dynamics of ideas spread, innovations and computer viruses [46, 47].

A result of particular interest as to the spread of the diseases and how this process relates to the network structure is provided by the research conducted by Pastor-Satorras and Vespignagni who studied the disease diffusion dynamics on random and scale-free networks. Their research showed that while in random networks a local infection contaminates the whole network only when the spreading rate exceeds a threshold limit, in scale free networks this threshold is zero meaning that the entire population of the network will be infected [25, 48].

When modeling and studying complex systems such as social networks which comprise a large number of elements with all of their details, it is vital that the dynamics of each element and the relations among them be clarified. Network topology and nodes states often are coupled critically in such a way that the nodes behavior is constrained by the structure which at the same time is generated by the behavior of the nodes. This kind of interaction generates unpredicted behavior which is called emergence and cannot be explained at the level of elements [49].

5.1 Statistical Physics Models

Valuable insights into the analysis of the behavior of complex networks with applications to social ones come from the field of the statistical physics. This scientific area provides for a large number of tools catering to the prediction of a system behavior by observing the properties of its elements, e.g. how magnetism appears out of the collective behavior of millions of atoms and the regulating factor of temperature [20]. Based on this observation it could be interesting to investigate the potential existence

of temperature-like parameters which when tuned to a critical value could lead to the formation of leaderless or hierarchically structured communities [17]. Following such an approach in the modeling of social networks could facilitate the examination of the effects of external signals or perturbations on the topology and dynamics of social networks (network-environment interaction). In this modeling scenario it could be observed that a social system reacts macroscopically even to a microscopic external perturbation thus demonstrating a behavior characterizing critical phenomena in which a dissipative system tends to rearrange itself in a way that enables the development of long-range temporal and spatial correlations (sand pile model) [31]. While many studies explore the effect of a network's nodes interaction in relation to a network's evolution, the case of network-environment interaction is largely unexplored as to the peculiarities of information processing in self-organised critical networks and the idea of finest adaptation at the "edge-of-chaos" [28].

For the study of the opinion dynamics on an adaptive random network, models from statistical physics have been successfully used thus explaining phenomena such as opinion formation, voting preferences, information diffusion and spread of diseases. Phenomena like these include a large number of "agents" which evolve over time due to the existence of external influence and noise. The final states emerge as a result of nonlinear dynamics, critical phenomena and phase transitions the outcome of which depends on the initial states and driving forces [32].

5.2 Evolutionary Models

Evolutionary models can be particularly helpful in examining the mechanisms underlying the growth of social networks. Bak-Sneppen model for instance, explores the feedback mechanism between fitness dynamics and topological restructuring. According to this model the node with the lowest barrier is the first to evolve as a result of interactions with the neighboring nodes. However the mutation of one node changes the state of all the interacting nodes thus affecting their fitness. When the aforementioned procedure is repeated the social network self-organizes itself in a critical stationary state in which all the barriers are distributed over certain threshold value. Given that socio-economic systems display a strongly networked structure the foregoing model could be used for examining the interactions between socio-economic nodes be they individuals, firms or trading countries [17].

5.3 Threshold Models

Valuable insights into how diffusion takes place in online networks is found in the in the work of Goel et al. [50]. Their paper discusses how Granovetter [51] and Lopez et al. [52] have modeled adoption decisions –in particular the expensive ones – using the "threshold" concept whereby adoption occurs only after a number of individuals (the threshold value) have adopted. Dodds and Watts [53] have proposed a model of "generalized contagion" that combines disease spread dynamics and threshold models. Young introduced a model of observational learning which resembles the threshold behavior [54]. These models demonstrate that the local and global structure of a

social network can substantially affect the size and the probability of cascades initiated by any given seed. Many adoption models have shown that a cascade triggered by a single seed can cover the entire network provided that the circumstances are proper - as for example in forest fires which in order to be large require the proper combination of favorable factors including wind, temperature, humidity and inflammability of trees. By the same token, cascades of social influence require the appropriate combination of many factors in order to spread throughout the network thus indicating the chaotic nature of the spread dynamics, since small differences in the initial conditions could have disproportionately large outcomes. However, one important aspect that has to be taken into consideration when studying the social influence dynamics, is that the social network nodes might be highly inhomogeneous in terms of their inherent thresholds. Nodes with high thresholds are slow and switch their state only for a limited set of input configurations (similar to the effect of canalizing functions in random boolean networks). On the contrary nodes with low thresholds are more likely to switch [28].

5.4 Neural Networks Models

The identification and modeling of the mechanisms that govern the topological development of social networks could be enabled by introducing approaches applied in the neural networks governed by the Hebbian rule suggesting that links are reinforced between neurons being at similar states [34]. Gross' and Blasius' [30] survey regarding adaptive coevolutionary networks, provides two different approaches in studying them. The first focuses on the values of important topological properties of evolving networks and the second examines how the functioning of the network hinges on these properties.

5.5 Cellular Automata Models

Another approach to studying the dynamics of adaptive social networks is the use of cellular automata [49]. In these models cells with discrete states are placed on lattice topology and their state transition is controlled by the states of its neighbors. Sayama and Laramee in their survey of modeling methods for complex systems present the "Generative Network Automata" - a novel modeling technique which integrates the nodes states transitions and the network topology transformations of complex systems [55].

5.6 Adaptive Voter Models

Opinion formation dynamics developed on social networks could be studied by using adaptive voter models where the network topology coevolves with the state of the nodes. This model despite its simplicity can also be effective when applied to more complex systems in which a number of different opinions interact by a pairwise-symmetric competition [56]. Benczik et al [32], in their study about opinion dynamics on adaptive networks propose Voter and Voterlike models for analysing phenomena

stemming from human behavior such as the emergence of collective organisation in socio-cultural situations. Voter model has also broad applications in the sociophysics field where in combination with statistical physics tools like mean-field approaches or numerical simulations has been broadly used for studying the spreading of cultures, religions, languages or political opinions. The Axelrod and Sznajd [57] models are particularly developed for studying the specific types of opinion dynamics.

5.7 Molecular Network Models

A new type of analysis applied in social networks comes from the field of biology and especially the analysis of molecular networks. This kind of analysis focuses on the discovery of local patterns in a network, called motifs [26]. Motifs are actually over-represented small sub-graphs existing in a network. Similarly activity motifs are over-represented patterns in the attributes of nodes and network edges.

5.8 Agent-Based Models

Social dynamics can also be studied by means of agent-based models which can simulate how the individual behavior leads to global transformations of network structure [13, 58, 57]. In this modeling technique, the actions of the agents which can be either individuals or groups following specific behavioral rules, are intertwined thereby affecting the overall network since the actions of each agent have a defining impact on the actions of the others. In this respect, changes in the network structure can be predicted on the basis of prior knowledge of the agents actions rules [14].

5.9 Social Percolation Models

With social adaptive networks encompassing the properties of “small world”, “high clustering coefficient” and “power law” degree distribution [20, 59, 60], the study of network dynamics should be combined with the topological network connectivity features which can give rise to emergent phenomena of particular interest. A theoretical framework with the ability to extract global network properties from the local specifications is that of *percolation* which as theory lies at the crossroads of probability theory and topology [61]. A system percolates when an adequate part of its constituent elements is connected locally so that a global connection emerges.

By applying percolation theory to marketing we could explain why some products succeed and other fail. According to percolation theory if globally connected clusters do not emerge as the information about the product passes among consumers, then the commercial success of the product will be limited as there will be consumer clusters which will never be exposed to the product and as a result will never adopt it [62].

Percolation models are suitable tools for studying the emergence of phase transitions, that is a sudden and striking change in the qualitative behavior of a system triggered by a subtle change in the initial conditions. Percolation models could also be used for predicting whether a large-scale diffusion will take place or not as the

propagation may vanish before a crucial fraction of the system is reached by the diffusion dynamics [62].

The shift to the percolation regime is usually sudden and the values of the control parameters at which the transition takes place are called "critical values". For example, adjusting this control parameters over and below the critical values can make the difference between the commercial failure and success of a product. Mort [63] proposed the application of percolation theory to marketing. Solomon and Weisbuch [64] cast the diffusion of information, beliefs, products and behavior in percolation terms thus introducing the term *"Social Percolation"*. The *"Social Percolation"* framework shapes the internal social forces as local interactions between neighboring nodes and monitors the resulting adoption patterns.

5.10 Summary of CAS Models for Analyzing Social Dynamics

The foregoing approaches originating from the Complex Adaptive Systems theory, are summarized in table 2 associating types of CAS models and the social phenomena each of them is mostly suitable for analyzing.

Table 2. CAS models for analyzing social phenomena

Complex Adaptive System Approaches for Analyzing Social Dynamics	
Models	**Social Phenomena**
Statistical Physics Models	Opinion Formation, Voting Preferences, Information Diffusion, Spread of Diseases
Evolutionary Models	Growth of Social Networks
Threshold Models	Adoption Decisions
Neural Network Models	Topology Changes in Social Networks
Cellular Automata	State Transitions of nodes
Adaptive Voter Models	Opinion Formation Dynamics, Collective Organization in Socio-Cultural Situations, Spreading of Cultures, Religions, Languages, Political Opinions
Molecular Network Models	Activity Motifs (over-represented patterns of behavior)
Agent-Based Models	Global Social Transformations resulting from Individual Behavior
Social Percolation Models	Diffusion of Information, Beliefs, Products, Behavior

6 Conclusion

We live in a world of networks. Communication networks, the internet, protein networks, chemical reaction networks, transportation networks, power grids, neural networks and social networks are some typical examples. However, although networks are ubiquitous their properties, structures and dynamics are not fully understood, thereby rendering the prediction of their behavior an intriguing and challenging research endeavor.

The increasing size of social networks coupled with the perpetual contact introduced by the omnipresent social networking tools, renders the social dynamics ever more complex and unpredictable. Finding the forces that drive the social transformations in a context of an overwhelmingly complex world affected by economic crises, business competition, societal disasters and security threats is a step towards gaining predictability power. To this end the study of online social networks from the viewpoint of complex adaptive systems can provide significant insights concerning social network dynamics. The methods and models presented in this paper can substantially contribute to the application of complexity theory in the study of social phenomena. Having answers as to how users build, grow, merge and break up communities, how influence is spread, how trust is built and interaction is transmitted, we will be able to explain why local events can potentially be global events with wider socioeconomic impacts.

References

1. Moreno, J.L.: Sociogram and sociomatrix. Sociometry 9, 348–349 (1946)
2. Wasserman, S., Faust, K.: Social network analysis: Methods and applications. Cambridge University Press (1994)
3. Wasserman, S., Galaskiewicz, J.: Advances in social network analysis: Research in the social and behavioral sciences. Sage Publications, Incorporated (1994)
4. Butts, C.T.: Social network analysis: A methodological introduction. Asian Journal of Social Psychology 11, 13–41 (2008)
5. Bavelas, A., Barrett, D.: An experimental approach to organizational communication. American Management Association (1951)
6. Brass, D.J.: Power in organizations: A social network perspective. Research in Politics and Society 4, 295–323 (1992)
7. Burt, R.S.: V The Social Structure of Competition (1992)
8. Willer, D.: Network exchange theory. Praeger Publishers (1999)
9. Knoke, D., Kuklinski, J.H.: Network analysis: basic concepts. In: Thompson, G., et al. (organizadores) Markets, Hierarchies and Networks, pp. 173–182. Sage Publications, London (1991)
10. Burt, R.S.: Toward a structural theory of action: network models of social Structure, Perception, and Action (1982)
11. Scott, J.: Social network analysis: A handbook. Sage Publications Limited (2000)
12. Newman, M.E.J.: The structure and function of complex networks. SIAM Review 45, 167–256 (2003)

13. Hui, C., Goldberg, M., Magdon-Ismail, M., Wallace, W.A.: Simulating the diffu-sion of information: An agent-based modeling approach. International Journal of Agent Technologies and Systems (IJATS) 2, 31–46 (2010)
14. Scott, J.: Social network analysis: developments, advances, and prospects. Social Network Analysis and Mining 1, 21–26 (2011)
15. Wasserman, S., Pattison, P.: Logit models and logistic regressions for social networks: I. An Introduction to Markov Graphs and p. Psychometrika 61, 401–425 (1996)
16. Pattison, P., Wasserman, S.: Logit models and logistic regressions for social networks: II. Multivariate Relations. British Journal of Mathematical and Statistical Psychology 52, 169–193 (1999)
17. Caldarelli, G., Garlaschelli, D.: Self-organization and complex networks. Adaptive Networks, 107–135 (2009)
18. Watts, D., Strogatz, S.: The small world problem. Collective Dynamics of Small-World Networks 393, 440–442 (1998)
19. Watts, D.J.: Networks, dynamics, and the small-world phenomenon 1. American Journal of Sociology 105, 493–527 (1999)
20. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Reviews of Modern Physics 74, 47 (2002)
21. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science 286, 509–512 (1999)
22. Stauffer, D., Aharony, A.: Introduction to percolation theory. CRC (1994)
23. Newman, M.E.J., Watts, D.J.: Scaling and percolation in the small-world network model. Physical Review E 60, 7332 (1999)
24. Pastor-Satorras, R., Vespignani, A.: Epidemics and immunization in scale-free networks. arXiv preprint cond-mat/0205260 (2002)
25. Pastor-Satorras, R., Vespignani, A.: Epidemic dynamics in finite size scale-free networks. Physical Review E 65, 035108 (2002)
26. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of networks. Advances in Physics 51, 1079–1187 (2002)
27. Caldarelli, G., Vespignani, A.: Large Scale Structure and Dynamics of Complex Networks: From Information Technology to Finance and Natural Science (Complex Systems and Interdisciplinary Science) (2007)
28. Rohlf, T., Bornholdt, S.: Self-organized criticality and adaptation in discrete dynamical networks. Adaptive Networks, 73–106 (2009)
29. Murray, G.M.: The quark and the jaguar: Adventures in the simple and the complex. Published by Little, Brown and Company (UK) Limited, London (1994)
30. Gross, T., Blasius, B.: Adaptive coevolutionary networks: a review. Journal of the Royal Society Interface 5, 259–271 (2008)
31. Miller, J.H., Page, S.E.: Complex adaptive systems: An introduction to computa-tional models of social life. Princeton University Press (2007)
32. Benczik, I.J., Benczik, S.Z., Schmittmann, B., Zia, R.K.P.: Opinion dynamics on an adaptive random network. Physical Review E 79, 046104 (2009)
33. Traulsen, A., Santos, F., Pacheco, J.: Evolutionary games in self-organizing populations. Adaptive Networks, 253–267 (2009)
34. Hebb, D.O.: The organization of behavior. Wiely, New York (1949, 2002)
35. Do, A.L., Gross, T.: Contact processes and moment closure on adaptive networks. Adaptive Networks, 191–208 (2009)
36. Holme, P., Newman, M.E.J.: Nonequilibrium phase transition in the coevolution of networks and opinions. Physical Review E 74, 056108 (2006)

37. Kauffman, S.A.: The origins of order: Self-organization and selection in evolution. Oxford University Press, USA (1993)
38. Shaw, L., Schwartz, I.: Noise induced dynamics in adaptive networks with applications to epidemiology. Adaptive Networks, 209–227 (2009)
39. Barabási, B.Y.A.L., Bonabeau, E.: Scale-Free. Scientific American (2003)
40. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the internet topology. ACM SIGCOMM Computer Communication Review, 251–262 (1999)
41. Mislove, A., Marcon, M., Gummadi, K.P., Druschel, P., Bhattacharjee, B.: Measurement and analysis of online social networks. In: Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, pp. 29–42 (2007)
42. Catanese, S., Meo, P., Ferrara, E., Fiumara, G., Provetti, A.: Extraction and analysis of facebook friendship relations. Computational Social Networks, 291–324 (2012)
43. Ghoshal, G.: Structural and dynamical properties of complex networks (2009), http://141.213.232.243/handle/2027.42/64757
44. Palla, G., Pollner, P., Barabási, A.L., Vicsek, T.: Social group dynamics in networks. Adaptive Networks, 11–38 (2009)
45. Gross, T., Sayama, H.: Adaptive Networks, pp. 1–8 (2009)
46. Coleman, J., Katz, E., Menzel, H.: The diffusion of an innovation among physicians, Sociometry 253–270 (1957)
47. Valente, T.W., Davis, R.L.: Accelerating the diffusion of innovations using opinion leaders. The Annals of the American Academy of Political and Social Science 566, 55–67 (1999)
48. Pastor-Satorras, R., Vespignani, A.: Epidemic spreading in scale-free networks. Physical Review Letters 86, 3200–3203 (2001)
49. Tomita, K., Kurokawa, H., Murata, S.: Graph-rewriting automata as a natural extension of cellular automata. Adaptive Networks, 291–309 (2009)
50. Goel, S., Watts, D.J., Goldstein, D.G.: The structure of online diffusion networks. In: Proceedings of the 13th ACM Conference on Electronic Commerce, pp. 623–638 (2012)
51. Granovetter, M., Soong, R.: Threshold models of diffusion and collective behavior. Journal of Mathematical Sociology 9, 165–179 (1983)
52. Lopez-Pintado, D., Watts, D.J.: Social influence, binary decisions and collective dynamics. Rationality and Society 20, 399–443 (2008)
53. Watts, D.J., Dodds, P.S.: Influentials, networks, and public opinion formation. Journal of Consumer Research 34, 441–458 (2007)
54. Young, K.S.: Internet addiction: The emergence of a new clinical disorder. Cyber Psychology & Behavior 1, 237–244 (1998)
55. Sayama, H., Laramee, C.: Generative network automata: A generalized frame-work for modeling adaptive network dynamics using graph rewritings. Adaptive Networks, 311–332 (2009)
56. Demirel, G., Prizak, R., Reddy, P.N., Gross, T.: Opinion formation and cyclic dominance in adaptive networks. arXiv preprint arXiv:1011.1124 (2010)
57. Axelrod, R.: The complexity of cooperation: Agent-based models of competition and collaboration. Princeton University Press (1997)
58. Helbing, D., Balietti, S.: How to do agent-based simulations in the future: From modeling social mechanisms to emergent phenomena and interactive systems design. Tech. Rep. 11-06-024, Santa Fe Institute, NM, USA (June 2011), santa Fe Working Paper (2011)
59. Pastor-Satorras, R., Rubí, M., Diaz-Guilera, A.: Statistical mechanics of complex networks. Springer (2003)

60. Newman, M., Barabasi, A.L., Watts, D.J.: The structure and dynamics of networks. Princeton University Press (2011)
61. Hunt, A., Ewing, R.: Percolation Theory: Topology and Structure. Percolation Theory for Flow in Porous Media, 1–36 (2009)
62. Erez, T., Moldovan, S., Solomon, S.: Social Anti-Percolation and Negative Word of Mouth. arXiv preprint cond-mat/0406695 (2004)
63. Mort, J.: Perspective: the applicability of percolation theory to innovation. Journal of Product Innovation Management 8, 32–38 (1991)
64. Solomon, S., Weisbuch, G., de Arcangelis, L., Jan, N., Stauffer, D.: Social percolation models. Physica A: Statistical Mechanics and its Applications 277, 239–247 (2000)

Evaluation of Spectrum Rights for Radio Broadcasting Operators

Spyros E. Polykalas[1,*] and Ulrich Stumpf[2]

[1] TEI of the Ionian Islands Lixouri, Kefalonia, Greece
s.polykalas@teiion.gr
[2] WIK Bad Honnef Germany
u.stumpf@wik.org

Abstract. The paper explores the valuation of radio spectrum granted by the State to radio operators for radio program broadcasting. Radio spectrum may be one of the most tightly regulated resources of all time, since every wireless device, from mobile phones to traditional radio devices, depends on access to the radio frequency wireless spectrum. Due to the scarcity of radio spectrum, the available number of radio broadcasting licenses is limited, therefore the State usually offers the licenses through open tender procedures following several tendering options such as auctions and beauty contests. One of the most crucial factor in the licensing procedures, is the valuation of the offered spectrum. The valuation of the spectrum determines the amount of money that radio operators should pay to the State, in order to hold the spectrum license for a fixed period of time. In this paper a cost model for radio spectrum evaluation is applied in order to identify the most crucial parameters that affect the determination of the spectrum price. The dynamic structure of the model allows a sensitivity analysis study in relation to the assessment of the variation effect of several parameters to the determined spectrum price. The appropriate determination of the examined parameters allow the State to maximize the government revenues but most important to grant the spectrum with the most efficient way.

Keywords: Spectrum Policy, Evaluation of Spectrum Rights, Radio Broadcasting, Spectrum value.

1 Introduction

Traditional radio stations broadcast theirs programs through frequencies allocated for radio broadcasting service. Each radio station holds a license in order to be able to broadcast its program through the allocated radio frequencies. Due to scarce characteristics of radio spectrum the available number of radio licenses is limited. Therefore radio stations should apply to State in order to grant a radio license. The radio stations operate under a license for a fixed period. Once this license expires, the State either initiates an open allocation procedure such us an auction or a beauty contest, or extent the existing licenses, if the latter is foreseen in the relevant legal

* Corresponding author.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 141–152, 2013.

framework. In both cases, the State should set the appropriate financial terms in order to promote spectrum efficiency, while in parallel to maximize the public revenues for the State.

The new EU framework for electronic communication sets the principles for radio frequency licensing and usage, including the principles should be adopted in the case of frequencies used for radio broadcasting. In particular in article 5 of the Authorization Directive (2002/20 Directive as amended in 2009) is foreseen that Members States shall ensure that the rights of use for radio frequencies shall be granted through open, objective, transparent, non-discriminatory and proportionate procedures. In addition in the same article is mentioned that the Member States shall ensure, that radio frequencies are efficiently and effectively used in accordance with the provisions of Framework Directive (2002/21 Directive as amended in 2009) in relation to competition promotion and harmonization of use of radio frequency. Although the current financial crisis forces the States to find ways to maximize the revenues in each licensing procedure the relevant EU Framework, clearly promote the efficiency usage of radio frequency, the competition and the provision of innovative services. Indeed as stated in the point 32 of the preamble of the Authorization Directive the imposed fees should not hinder the development of innovative services and the competition in the relevant market.

The efficient assignment of the spectrum and in general the efficient spectrum management has been several times subject for discussion both from academic and industry point of view. In particular in [4] a techno-economic analysis was conducted in order to examine the impact of the different spectrum management approaches to the social welfare. The authors concluded that regimes allowing secondary use, whether by sharing or by trading, have greater preferences, in most cases, than other approaches. In addition in [5] the degree of substitutability of frequencies with or without regulatory constraints was examined. The authors concluded that the regulatory constrains are a major source of limitations on substitutability. In addition alternative forms of intervention were considered, including caps on spectrum holdings or on the acquisition of spectrum at any award. In [6] the authors examined the policy trends towards more flexible forms of spectrum management. In [7] the authors argued that when the spectrum is allocated to operators and high investments are required, the government should consider real options methodology for setting the price of the license, or the base price of the licenses in case the government decides to follow an auction methodology to allocate the licenses to determine a more accurate price of the license which takes into account the managerial flexibility.

In [8] the authors argued that the Discounted Cash Flow (DCF) valuation is the most common method to value real assets whose future cash flows can be forecasted with certain degree of predictability. They believed that if a project's risk differs considerably from the firm's average risk then the WACC is adjusted upwards or downwards to arrive at the new discount rate for the project depending on whether the project is more risky or less risky respectively. They criticized DCF method for one of its inherent and structural weakness which is that the project's value will remain same and unaffected despite any future decisions by the management of the firm or the project.

In [9] a methodology is proposed for the calculation of licenses fees for radio broadcasting, based on an objective assessment of the value of the spectrum for an averagely efficient entrant. It is argued that as incumbents have made specific investments to operate a license and have an installed base of listeners, they can be expected to have a higher valuation of the spectrum than an entrant. The value of the spectrum for an averagely efficient entrant would be the second highest bid and therefore the expected price that, the incumbents would have to pay for renewing their license in case a (second bid) auction were held.

Compared to existing research efforts, this paper is more closed to [9] but its innovation could be summarized in the following three reasons: (a) it proposes a fully dynamic model for spectrum evaluation used for radio broadcasting (b) it uses data from a real market but could be easily applied to any other market ,(c) it identifies the most crucial parameters that affect the determination of the spectrum price and (d) its conclusions could be used by policy makers for efficient spectrum allocation, promoting the competition and the provision of innovative services.

The rest of the paper is organized as follows. Section 2 expands on the spectrum evaluation available options. Section 3 presents in details the model used for the spectrum evaluation, while in Section 4 the results of the model are discussed, attempting the determination of the most crucial licensing factors. The conclusions of our study are presented on Section 5.

2 Spectrum Valuation

For the evaluation of radio spectrum by the State authorities, two methodologies are mainly used, the benchmarking approach and the modelling approach. Indeed in [10] the authors concluded that in Europe radio broadcasting licenses are usually awarded by means of a beauty contest or, less frequently, an auction.

In the benchmarking approach, prices/costs from other countries are used with appropriate adjustments. The disadvantage of this method is that the prices in each market depend on national circumstances such as the size of the offered spectrum, the used technologies, the legal framework, the local economy the level of competition, as well as the projections for the future revenues growth and costs. The adjustment of all these factors into the implied values requires full knowledge of the national circumstances in each case and of the related factors under examination.

The most used methodology in the modeling approach is the Discounted Cash Flow (DCF), since the radio spectrum is an asset that is not trade on a commercial market, therefore an operator interested to buy a radio spectrum right is willing to pay based on what he could earn using the radio spectrum. In the DCF the future costs and revenues of a spectrum operator are estimated, in order to determine the relevant Net Present Value (NPV). It could be said that the NPV represents the maximum price that an operator would be prepared to pay for a given amount of spectrum, since any higher payment would not lead to a reasonable return on its investments for a given duration of the license. It should be also noted that, in some cases the DCF methodology may result in uncertain estimates, since the future cash flows are based on projections of the spectrum operator's revenues and costs. In addition, the longer the license duration, the higher the uncertainty of the projections, resulting in

uncertain estimates for the NPV. Nevertheless the DCF methodology is accepted as an appropriate methodology for spectrum valuation.

3 DCF Modelling Approach

The model has been implemented in Microsoft Office Excel using Microsoft Visual Basic. The DCF model, by determining the future cost and revenues of a radio broadcasting operator, provides the required inputs in order to set appropriate radio spectrum fees. The accuracy of the model outputs depends, among others, on the extent to which the DCF model reflects all the relevant current and future revenues and costs of a modelled radio operator. To do so several assumptions, reflecting actual market condition, should been adopted.

The user of the model is able, via the Graphical User Interface (GUI) of the model, to examine several scenarios by changing a variety of model input parameters. As results, the model provides an estimation of the NPV (Net Present Value) of the discounted cash flows for a potential license owner, during the lifetime of the radio license, as well as the relevant Terminal Value (TV). The TV indicates mainly the value of the radio operator's infrastructure at the end of the license period. The NPV can be used as an indicator, for the determination of the spectrum fee that should be charged in the relevant licensing procedure.

Figure 1 shows the structure of the DCF model. For reasons of simplicity, only the high level architecture of the model is presented.

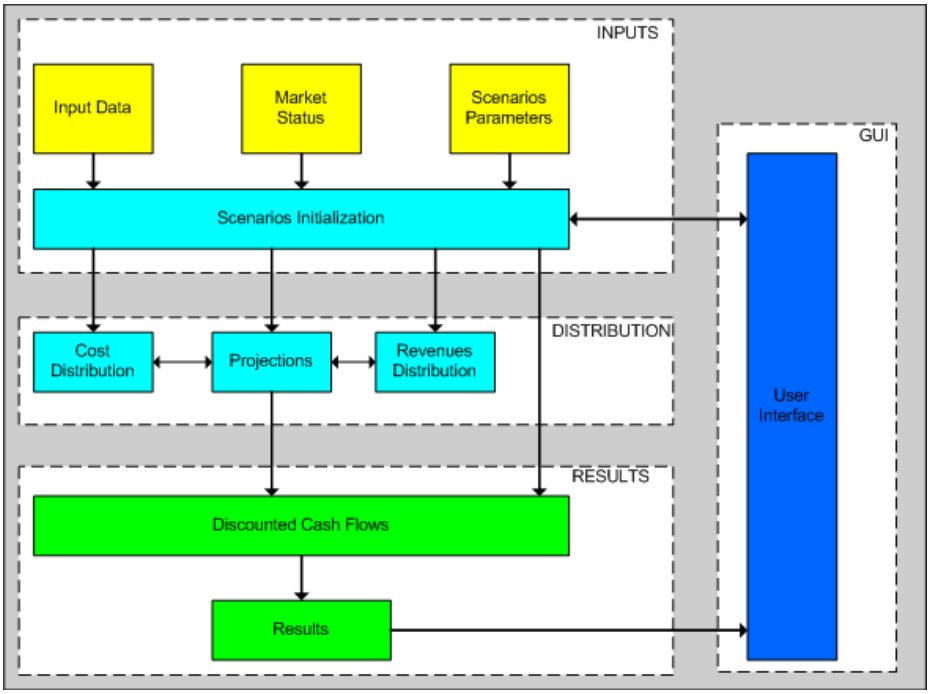


Fig. 1. Structure of the DCF model

As depicted in the above Figure the model consists of four main segments: the Inputs, the Distribution, the Results and the Graphical User Interface. In the following paragraphs is described in detail each of the model segments.

The Inputs Segment contains the inputs required for the determination of the parameters of each scenario. It consists of four sub-segments: Input Data, Market Status, Scenarios Parameters and Scenarios Initialization. The input data are grouped in the following three data categories: Revenues, OPEX and Investments. Revenues data are related to revenues coming from advertising as well as to all other revenues from operations related to radio broadcasting. Operating expenditures (OPEX) are related to the operating costs for the provision of a radio broadcasting service. Investments are related, to the investments made in the past, as well as, to the future investments in relation to radio broadcasting operation. Market Status sub-segment contains the parameters which are related to national conditions in the radio broadcasting market. In particular it contains data related to required network infrastructure for each type of geographical license (for example local, regional and national), such as the number of required sites and the number of required transmission links. It should be noted, that as regards the local licenses, the model provides two categories, one for urban areas and one for rural areas. The user of the model - through the GUI - is able to determine the number of areas that should be considered as rural and urban, taking into account the local market conditions such as the population, the number of radio listeners and the revenues of local radio operators. Scenario Parameters sub-segment contains all the rest parameters required for the initialization of each examined scenario. The fourth sub-segment, Scenarios Initialization, initializes all the required parameters for the running of each scenario. To do so, it receives inputs from the other three sub-segments of Inputs Segment and interacts through the GUI with the user of the model.

The Distribution Segment consists of three sub-segments: Projections, Cost Distribution and Revenues Distribution. The Cost Distribution allocates the existing and future costs related to radio broadcasting operation over the duration of license. In particular the costs are allocated as depreciation expenses to each period in which the asset is used, beginning when the asset is placed for the provision of radio broadcasting service. The level of existing infrastructure and future investments are mainly driven by the geographical type of license (national, regional, local), as well as whether in the examined scenario the radio operator is an existing one or a new entrant. For the allocation of each type of cost, an estimation of the useful lifetime of each asset has been taken into account, while the straight line (linear) depreciation approach has been chosen. In straight-line depreciation, a yearly portion of original cost in equal increments is allocated over the lifetime time of each asset. It should be noted that the salvage value of the assets has been ignored. The lifetime of each group of assets is determined by the user of the model through the GUI of the model. The Revenues Distribution determines the revenues of a radio broadcaster on a yearly basis, during the period of the license. The future revenues and costs (OPEX and CAPEX) are driven by the Projection sub–segment parameters, where the user of the model is able to examine several scenarios. In particular for both revenues and opex

costs, several options for linear increase or decrease are available. For future investments two options are available: in the first option, it is assumed that the future investments of the radio operators, during the period of the license, are based on the operator's relevant projections for the years 2012, 2013 and 2014, while in the second option the future investment are determined as a portion of the future revenues of radio broadcasting operators. The level of portion of the future revenues allocated to investments is determined by the user of the model. The revenues of each radio operator are driven by the relevant market share. In our scenarios, equal market shares have been considered between radio operators for each geographical type of license. Given that the DCF analysis will be truncated after the duration of the license, it is considered appropriate to take into account a terminal value (remaining value) that reflects the continuity of the business and the on-going value of the business's assets. Thus the non-amortised values of the assets at the end of the license period has been calculated as the remaining value (or terminal value) of a radio operator at the end of the license period. It should be noted that only the assets related to infrastructure have been taken in to account for the estimation of the terminal value.

The Results Segment contains the outputs of the DCF model and consists of two sub-segments: Discounted Cash Flows and Results.The Discounted Cash Flows interacts with the relevant sub-segments of Distribution and Inputs segments. The discounted cash flows are estimated taking into account the future cash flows, the discount rate (WACC) and the terminal value at the end of the license period. The formula of the discounted cash flows is:

$$DCF_y = \frac{\left(R_y - C_y - T_y\right) - I_y - CWC_y}{(1+i)^y} \tag{1}$$

with

$$T_y = t(R_y - C_y - D_y) \tag{2}$$

where,

DCF: Discounted cash flow,

R: Revenues,

C: Operational costs (distribution and non-distribution),

T: Taxes,

I: Investment expenditures (distribution and non-distribution),

D: Depreciation (distribution and non-distribution),

i: discount rate (cost of capital - WACC),

y: Year,

t: taxes,

CWC: Change in Working Capital (represents the operating liquidity of an operator),

The NPV is derived from the following formula:

$$NPV = IV_0 + \sum_{1}^{n} DCF_y + TV_{n+1} \tag{3}$$

where,

IV: Initial value (distribution and non-distribution),
TV: Terminal value (distribution and non-distribution),
n: License period.

The Results sub-segment contains the outcome of the model which is the NPV, and the TV of each scenario. As discussed previously, the Terminal Value represents the value of a radio broadcasting operator at the end of the licensed period. Since the TV mainly reflects the non-amortized values of the assets, and in particular, the assets related to network infrastructure, it is considered appropriate to represent, in each scenario, the relation between TV and NPV. Depending on the extent to which the non-amortized assets at the end of the licensed period could be re-used in other operations, not necessarily in a radio broadcasting operation, the total or a portion of the TV should be taken into account for the determination of the relevant spectrum fee. It is also assumed that the Initial Value of the existing operators is zero, since the existing infrastructures have been taken into account for the estimation of the required future infrastructure per type of operator and license.

The Graphical User Interface (GUI) Segment provides the appropriate interface to the user of the model, in order to set the inputs parameters of each scenario, as well as, to initialize the variation of the input parameters. In addition, as results, the GUI segment provides the NPV and the TV of each scenario.

4 Model Results

The input values used in the model were taken from the Greek national market. Several sources have been used such as the Greek National Council for Radio and Television (NCRTV) [1] and other national public sources such as [2], [3]. When national data were not available public benchmarking input data have been used such as [4]. Although the majority of the input data, was taken from the Greek market, we argue that the results of our analysis can be easily adopted in any other national radio broadcasting market mainly for the following reasons: i) our conclusions are based on the absolute variance of the parameters focusing on the influence of each parameter to the outputs of the DCF model, ii) the business model of a radio broadcasting operator used in the DCF model, could be applied in any radio broadcasting operator outside of Greece and iii) the network architecture and network components of a radio broadcasting operator used in the DCF model, could be also applied in any radio broadcasting operator outside of Greece.

In each of the following graphs the vertical axis shows the NPV and TV in nominal basis while horizontal axis shows the variation of examined parameter. In each graph the NPV and TV actual values have been normalized to the NPV's actual value of the first value of the examined parameter. It should be mentioned here that our aim is to

reveal the sensitivity of each examined parameter to the NPV and TV and not to determine the actual values of the NPVs and TVs. In addition in each examined scenarios only one parameter varies (the examined parameter) while all the other parameters remain constant, in order to reveal the sensitivity of the NPV and TV to the examined parameter.

In the first set of scenarios the following policy parameters are examined: Number of licenses for each geographical type of license, Number of areas for each type of geographical area, License duration, and Type of operator (existing or new entrant).

In Figure 3 the impact of the variation of the number of available licenses for each geographical type of licenses to the NPV and TV is presented. As expected, the NPV and the TV are decreased, as the number of available license increases. The impact of number of license variation is higher, in the case of national license, since as shown in the following Figure, an increase of the available number of licenses from 4 to 6, leads from positive values of NPV to negative values. As regards the local urban licensing it seems that even for 90 licenses the business case is profitable. We should mention here that based on the Greek national market conditions, very high percentage of the national radio broadcasting revenues (mainly advertising revenues) has been allocated to the radio broadcasters operating in local urban areas. On the other hand it seems that there is no business case for local rural radio operators, when the number of licenses exceeds a specific threshold. In our case this threshold is determined on five but we should mention that this threshold is highly sensitive, among others, to the distribution of advertising revenues.

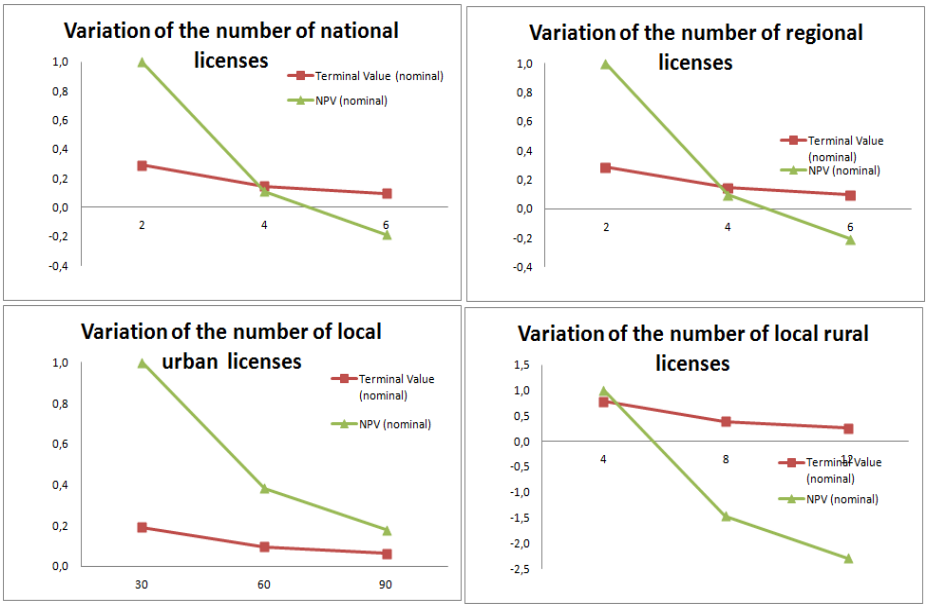


Fig. 2. Variation of the number of geographical license type

In the next set of scenarios the impact of the number of geographical areas per type of areas (regional, local urban and local rural) is examined. As presented in Figure 4, the NPV is more elasticity to number of areas variation than the TV. This is due to the fact that the TV mainly depends on the level of network infrastructure required for population and geographical coverage, while the NPV highly depends, among others, on the operators' revenues. As depicted in the following graphs, the variation of the number of areas has the higher impact on the NPV values in the case of local areas.

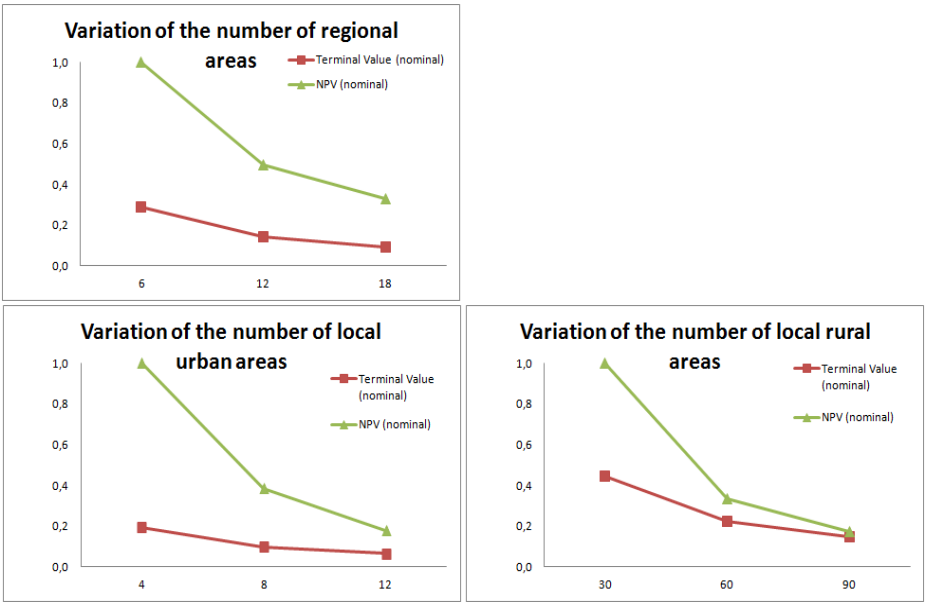


Fig. 3. Variation of the number of areas

In the left diagram of Figure 5 the impact of license duration in the NPV and TV for local urban license is presented. The national legal frameworks foresee license life time with specific duration (for example 6 years), while in the most cases there is the option to only once license extensions, therefore policy makers should be aware of the impact of the license duration in the NPV. As shown in the diagram below, an increase of the license duration leads to slight increase of the relevant NPV, while the TV is decreased. In the right diagram of Figure 5, the NPV and TV of an existing operator and new entrant are compared, both for national license radio operator. The NPV in the case of new entrant is almost 26% lower than the relevant NPV of an existing operator, due to the higher level of the required investments mainly on the network infrastructure.

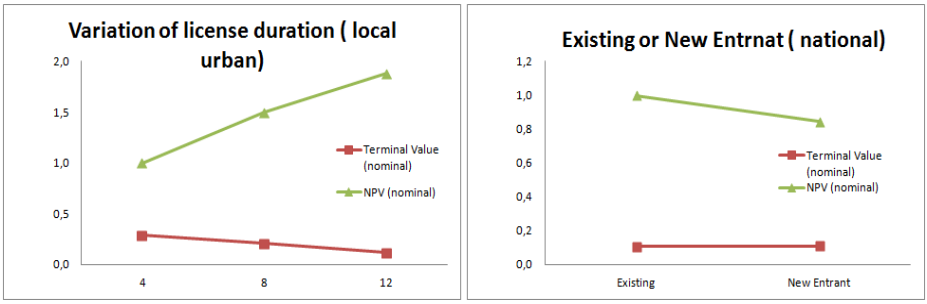


Fig. 4. Variation of License duration and Existing or New Entrant

The next set of scenarios deals with the future market conditions in relation to revenues projections. Due to the high uncertainty of future conditions, mainly due to financial crisis, the examination of the impact of revenues variation to the NPV should be also under consideration. In Figure 6 is examined, the impact of revenues projections to the values of NPV and TV for a local urban operator. In the left diagram, a linear increase of revenues has been assumed, during the lifetime of license with yearly growth of 1%, 3% and 5%. As expected the impact of the level of yearly growth has significant impact on the value of NPV. An annual growth of 4% leads to NPV almost 125 % higher than in the case of 2% annual growth. In the right diagram the impact of annual decrease in revenues is shown. Similarly with the left diagram, a revenues degradation leads to significant decrease of the NPV depending on the rate of revenues decrease. An annual decrease of 4% leads to almost 20% lower NPV compared with the NPV of 2% annual decrease.

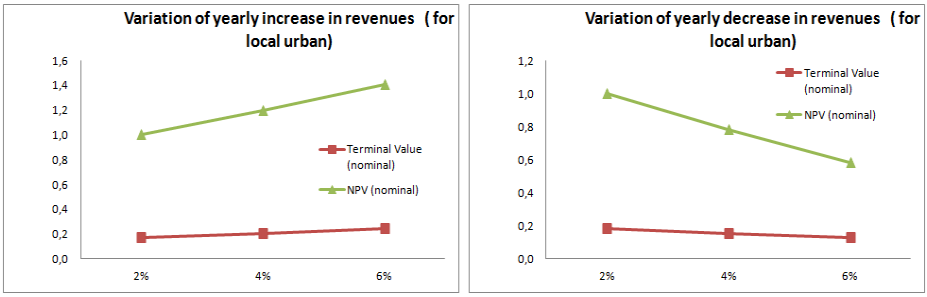


Fig. 5. Variation of revenues projections

In Figure 7 the impact of site' s opex variation to the NPV and TV of a national and local urban ratio operator is presented. Since the number of required sites in a case of national operator is a multiple of the relevant number in a case of local urban operator, the impact of the relevant site's opex is higher in the case of national licenses. Indeed an increase of 40% of the site's opex, decreases the NPV of a national operator by almost 11%, while the relevant NPV of a local urban operator remains almost constant (less than 4% decrease).

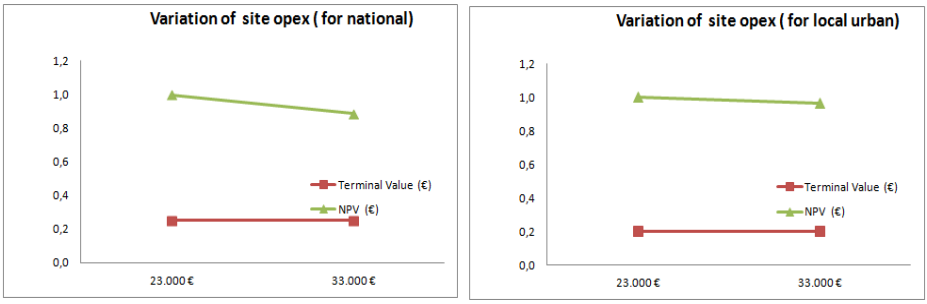


Fig. 6. Variation of site's opex

5 Conclusions

In the previous sections the valuation of radio spectrum granted by the State to radio broadcasting operators has been analysed. In particular a DCF model has been implemented for the affect examination of several parameters in the spectrum valuation. Since the spectrum valuation determines the amount of money that radio operators should pay to the State, in order to hold the spectrum license for a fixed period, an appropriate determination of the examined parameters will lead to efficient spectrum assignment.

The results of our analysis indicate the future revenues projections, the number of available licenses per geographical type of license, the number of areas and the duration of license as the most crucial parameters that affect the determination of the spectrum price.

As regards the number of licenses it is quite clear that depending on the geographical type of license an increase of the number of license may lead to a non profitable business case. The threshold which determines the profitability of a business model varies between the geographical type of license and depend mainly on the allocation of advertising revenues between the geographical types of licenses. In addition the allocation of advertising revenues between the types of licenses depends on national market conditions and may also varies between different national markets. In the issue of revenues projections the results of the model indicate that the spectrum valuation is highly sensitivity to revenues projections. Under current conditions, the projection of revenues, mainly advertising revenues, involves high uncertainty due to financial crisis, therefore policy makers should examine several scenarios in order to analyse future market conditions under different revenues assumptions. Last but not least the duration of a license plays an important role in the determination of the spectrum fees. In particular as the duration of licenses increases such the operators have more time to depreciate its investments and to increase its revenues therefore the determination of spectrum fees is directed connected with the license duration.

In addition the results of our model indicates that due to high sensitivity of spectrum valuation to a set of parameters, policy makers prior the determination of license procedures should examined in detailed each of these parameter taking utmost into account actual data from the national markets. The appropriate determination of

the examined parameters analysed in this paper, will allow policy makers not only to maximize the government revenues but also to grant the spectrum efficiently and effectively.

References

1. National Council for Radio and TV, http://www.esr.gr
2. Focus Bari market analysis company, http://www.focus.gr
3. Association of Companies for Marketing and Communications, http://www.edee.gr
4. Hwang, J., Yoon, H.: A mixed spectrum management framework for the future wireless service based on techno-economic analysis: The Korean spectrum policy study. Telecommunication Policy 33, 407–421 (2009)
5. Cave, M.: Anti-competitive behaviour in spectrum markets: Analysis and response. Telecommunication Policy 34, 251–261 (2010)
6. Delaere, S., Ballon, P.: The business model impact of flexible spectrum management and cognitive networks. Info. 9(5), 57–69 (2007)
7. Sinha, P., Mudgal, H.: Valuation of 3G Spectrum License in India: A Real Option Approach. Faculty of Management Studies, University of Delhi - Munich Personal RePEc Archive (2011)
8. Brandao, L.E., Dyer, J.S., Hahn, W.J.: Using Binomial Trees To Solve Real-Option Valuation Problems. Decision Analysis 2(2), 69–88 (2005)
9. Valuing Commercial Radio Licences, Study by SEO Economic Research (July 2011)
10. Study covering several European countries (Analysys Mason & Hogan & Hartson) (2009)

Using the NETC@RDS Approach as a Basis for Cross-Border Electronic Authentication

George Pangalos[3], Noel Nader[2], and Ioannis Pagkalos[1]

[1] Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki
[2] Direction des Programmes, International Projects Coordinator, GIE SESAM-Vitale, France
[3] Informatics Laboratory, General Dept. Faculty of Technology, University of Thessaloniki
{ipagkalo,pangalos}@auth.gr,
Noel.Nader@sesam-vitale.fr

Abstract. Many countries, European and worldwide, have increasingly issued during the last decade electronically readable identity documents to their citizens, for different purposes and applications. However, a major characteristic of all these systems is that they are basically available in a national context. For example, European citizens that move freely through the Member-States face the problem that their eIDs from their home state do not allow access to services of another Member-State in which they are temporarily present. Public Administrations are also unable to provide services to European citizens from other Member-States with the same ease and efficiency as they do to their national citizens. In order to avoid such confusing situations, cross-border services should be fully integrated in the national/regional and local information systems. It is, therefore, an important task to improve the cross-border interoperability of electronic identification and authentication systems. ENISA, the European agency for the security of computer systems and networks, recently published a report dealing with an important aspect of this problem: the security issues in cross-border electronic authentication. The report assesses the risks of electronic authentication in cross-border solutions and provides a generic implementation model. This paper describes an implementation methodology for addressing the cross-border interoperability of electronic authentication problem, based on the ENISA generic model. The proposed implementation methodology has been based on the successful NETC@RDS project approach and experience, described herein. This methodology can provide a suitable secure cross-border, multi-purpose authentication implementation based on the aforementioned generic model that can be used in various sectors.

Keywords: Electronic authentication, cross border authentication, security.

1 Introduction

Many countries increasingly offer citizens electronic access to their services [1,2,3]. These e-services often use electronic authentication (eID) and are usually implemented at a national level with specific technologies, specific security concepts and

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 153–168, 2013.

specific business logic. As a result, in most cases these systems can only be accessed from within the Member-State and by citizens of that state.

For example, during the last decade, several E.U. Member-States (France, Belgium, Germany, Austria, Italian regions, Slovenia and others) have distributed more than 200 million of health insurance smart cards to the population as evidence of entitlement for health care access and/or reimbursement at national level. A major characteristic of all these systems is that they are basically available in the national context. One reason for this is that citizen identification (like social and health insurance benefits) is usually related to services regulated at national or regional level. Existing solutions were, therefore, designed to be most efficient and fitting with respect to national requirements and infrastructures. Despite that, the goals of these systems are, in general, identical for all Member-State: managing identities, improving administrative efficiency, improving accessibility and user-friendliness, reducing abuse and fraud and reduction of costs.

Today, this may represent an undue restriction on the usage of these services to European citizens that move freely through all Member-States. This can lead to serious inequalities since, for example, health professionals might be reluctant to apply for cross-border benefits-in-kind procedures and, as a consequence, European patients might have to pay the bill for medical care delivered abroad themselves. Therefore, there is a need to extend these services beyond the national borders and beyond the user group of national citizens. At the same time, the European and national security and data protection laws and regulations must be respected and should not be undermined by any cross-border distribution of personal data.

Logically, the implementation of secure cross-border services, or the extension of a domestic system across borders, poses several challenges in the legal, organizational, security, semantic, socio-economic, and technical level.

An obvious technical challenge may be, for example, the fact that disparate IT systems with different technologies must be interfaced. Any problems rising from this are usually limited to designing a proper technical and financially affordable solution. Differences in the business logic of the national solutions are more difficult though. Health care and educational systems in particular differ greatly in the way that services are provided, evaluated and billed. Setting up a business or dealing with taxes is also very different from one State to another. In addition, amendments to the legal framework are often required in order to allow the distribution and processing of data by non-national institutions and organizations.

Another major prerequisite of any such e-government or e-health service is the trust in the authenticity of all participants and the provided data. Since most services of this type handle confidential data, the confidentiality must also be protected in a cross-border scenario. Some services also require a high availability if the citizen is not to suffer undue consequences. This establishes the need to discuss, evaluate and implement IT security in such cross-border applications.

ENISA, the European agency for the security of computer systems and networks, recently published an interesting report dealing with an important aspect of this

problem: the security issues in cross-border electronic authentication. More specifically, the report assesses the risks of electronic authentication in cross-border solutions and provides a generic implementation model [1].

This paper describes a possible implementation methodology for addressing this cross-border interoperability of electronic authentication problem, based on the ENISA report. The proposed implementation methodology has been based on the successful approach of the NETC@RDS[1] project and past experience in secure cross-border electronic authentication [2].

2 The ENISA Generic Models of Domestic and Cross-Border Electronic Authentication

According to the ENISA report terminology [1], in any domestic system that involves electronic authentication the User is assigned an electronic identity (eID). The scope of an eID may be limited to within the application (e.g. a health insurance number or a civil register). This eID is assigned by some entity within the system: the ID Authority. The ID Authority issues a Token (e.g. a health insurance card or a national ID card) to the User that identifies the user as the person with a specific eID. The token may contain the eID and other data in electronically readable form. The User Service Provider (e.g. a doctor or a vehicle registration office) interacts with the user and the user's token. He provides a service to the user that is linked to the application operated by the system operator. The laws, regulations or contracts governing the provision of this service require the user service provider to authenticate the user via the user's token against the system operator (figure 1).

The entire system, its participants, components and processes are governed by the same set of laws and regulations (Law A). These laws comprise the range from general regulations on the handling of personal data (e.g. based on the Data Protection Directive 95/46/EC [1]) to specific regulations regarding the application, the services or the token. While the range of possible (and existing) technical solutions and variations of tokens and electronic authentications is vast, the general principle is the same for all such systems. For example all such systems are homogenic with respect to technology, are governed by a single set of laws, and "know" all system participants, i.e. they are closed to non-participants.

When dealing with cross-border applications or utilizing a service with a token provided from outside the user service provider's state, the domestic model must be enhanced [1]. The notable difference to cross-border authentication as opposed to the domestic model is the fact that the User Service Provider (B) is actually the user service provider from another system, who is governed by different laws (Law B) and business rules. In addition, this other system may use different technology which may

[1] NETC@RDS service for the electronification of the European Health Insurance Card: a pan-European project supported by the EU's eTEN Program [2].

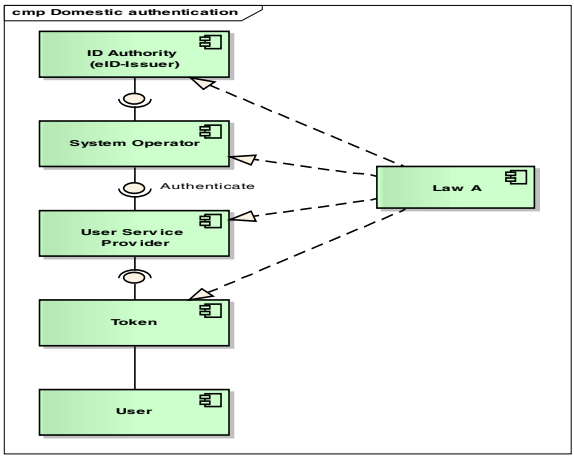


Fig. 1. The domestic ENISA electronic authentication scheme [1]

be incompatible. Even more important is the fact that the User Service Provider (B) is usually not known to the System Operator in the sense that there is often no direct contractual agreement and no clear-cut legal regulations that govern their relationship. Even worse, the laws governing the operations of User Service Provider (B) and System Operator are different, which raises all kinds of problems, from data protection to liability and insurance issues (figure 2).

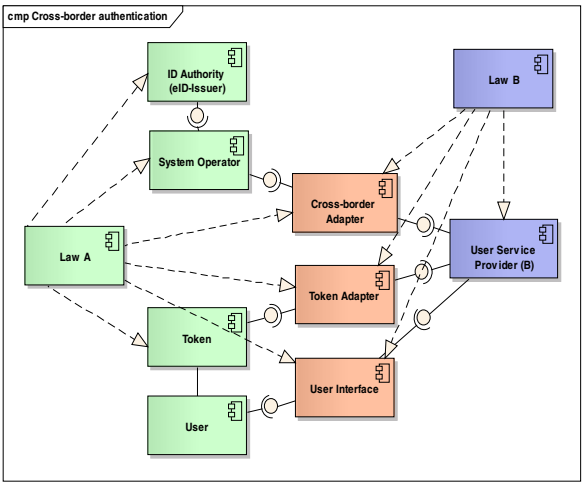


Fig. 2. The generic ENISA model of cross-border authentication [1]

In order to achieve compatibility of the two systems to the point that a user of the first system may receive services from the second system, two adapter components must be introduced into the systems [1].

The first one is the 'Cross-Border Adapter' which has the task of actually proxying an electronic authentication request from the local service provider (B) across the border between countries and systems to the system operator. This task includes the translation of data formats and business rules wherever necessary. The Cross-Border Adapter may be implemented in a number of ways. For each specific cross-border system the best and most appropriate implementation must be found. This is not so much a question of technology, but of possible solutions as defined by law and contractual agreements within the systems. The second component is the 'Token Adapter' which is specific to the cross-border solution of the system. Its main task is interfacing a token from one country with the user service provider from another country. Usually it may be considered to be an extension of the IT systems of the local service provider that is operated by him.

Comparing the ENISA generic model of a cross-border authentication system with the domestic system, some changes to the general principles of system design are evident which are relevant to any security evaluation. The cross-border system is heterogenic with respect to technology, is governed by two separate and at least partially disjoint sets of laws, and does not "know" all system participants, i.e. they are potentially open to non-participants.

Extending a domestic system to allow cross-border electronic authentication with a communication partner that is not native to the domestic system poses also a number of security challenges [1]. All of these challenges must be addressed and overcome to successfully implement cross-border interoperability. Examples include the different types of credentials that may link the user's identity to a token, the fact that the reliability of the credentials may differ, the wide range of different tokens used, the acceptance and trust of identity data coming from a foreign country, the authenticity check of a foreign token, and the authorization check of a foreign User Service Provider. Furthermore, the following important issues must also be taken into account:

Security Issues. International standards on evaluating information security and information security management systems can be found in various formats (e.g. the standards of the ISO 2700x family). However, these standards provide fundamental but often rather generic security requirements. According to the ENISA report, the core of any security evaluation (see also BSI 100-2 "IT-Grundschutz Methodology" [18]) is the definition of assets that must be protected and the protection requirements for these assets. Then each asset is assigned a protection requirement for the three basic protection values of confidentiality, integrity and availability.

Protection Requirements. The following assets are considered worthy of protection in the ENISA report [1]: Identity Data, Application Data, Token, IT systems of the system operator, User Service Provider, Cross-Border Adapter, Token Adapter and User Interface. It must be noted, however, that while they are discussed in the generic model, this can only be a starting point when it comes to the evaluation of a specific application. This is because the assessment of the protection requirements for each asset may differ greatly from application to application.

Other Technical and Legal Issues. Several other important technical and legal issues are also encountered when studying the ENISA generic model. For example [1]: the different types of credentials and their reliability, the tokens with different security levels that differ in their trustworthiness, the different technical infrastructures that elevate the amount of security vulnerabilities due to different security levels, the different authentication protocols and procedures that elevate the amount of security vulnerabilities due to different security levels and the attacks on the availability of the cross-border authentication process. Several legal problems also arise. For example, national restrictions on the transfer of identity data may differ, national regulations may prohibit authentication across borders, and the identity data may not be processed in an adequate, relevant and not excessive way to the purposes for which they are collected and/or further processed.

3 The NETC@RDS Approach for Implementing Cross-Border Electronic Authentication

3.1 The NETC@RDS Project

NETC@RDS is a pan-European project supported by the EU eTEN program which aims to improve the secure access of mobile European citizens to cross-border health care using advanced smart card and web services technologies [2,39]. More specifically, NETC@RDS aims to simplify health care access for citizens with health insurance evidence of entitlement while abroad and also to provide a reliable source of information for health care provider front office staff checking insured entitlement or initiating interstate billing/clearing procedures. It also aims to develop and use a Common Administrative Electronic Dataset for improved health insurance providers back office billing/clearing workflow applications and further modernization of post-processing activities.

NETC@RDS is basically addressing the following three business cases: (i) The automatic capture of the EHIC dataset, either by optically scanning the EHIC front layout or by electronically reading a national/regional health insurance smart card, (ii) The on-line verification of the EHIC dataset at the point of health care delivery against national/regional repositories located in the home country, and (iii) The sending the EHIC scanned copy, or the EHIC dataset to the competent institution, in view of further e-billing processing [2,39].

The NETC@RDS Consortium includes stakeholders from 16 European countries. It started in 2002 and ended in 2011. The service is however still provided today through the ENED consortium [39], created and supported by participating member states. The last implementation phase encompassed 626 health care service points in 16 EU/EEA Member states and Switzerland. The deployment of NETC@RDS infrastructure is also regarded as a test bed for the ongoing introduction of the e-EHIC.

3.2 System Overview

The NETC@RDS project [26,28] has established a cross-border online pan-European service to authenticate a patient's health insurance card and/or a patient's entitlement to health insurance benefits abroad for unplanned care. In the long run the overall goal of this project is the complete integration of the existing and future national infrastructures for health insurance claims in order to improve the data exchange.

The NETC@RDS technical architecture [2,26] consists of secure network interconnections within a Member state and between the Member states, linking national service portals and registries in each country with workstations within all service facilities. A cross-border mutual authentication is established every time a NETC@RDS user (typically a hospital clerk or a health practitioner) operates an online verification of the e-EHIC dataset as entitlement to receive health care abroad in one of the NETC@RDS service units/points. The NETC@RDS architecture currently features a direct communication between the individual national service portals as shown in figure 3.

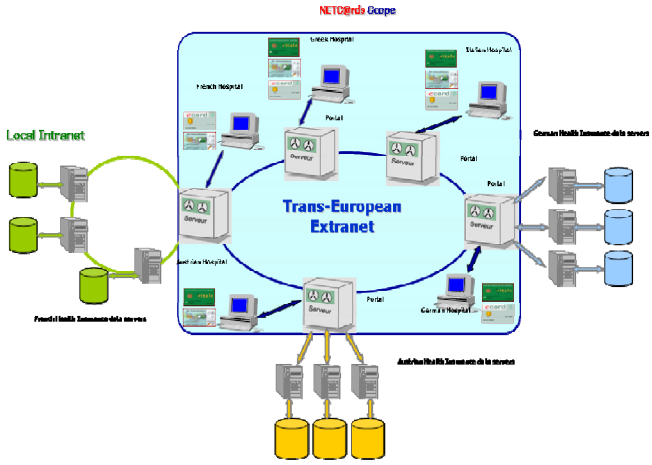


Fig. 3. How the existing NETC@RDS system works

3.3 *The NETC@RDS Secure* Cross-Border Electronic Authentication

Security is a critically important issue for NET@RDS deployment. Without adequate security in place none of the NETC@RDS systems can be used in real-life environments. The secure network interconnection between the 16 national portals relies on a common Information Security Systems Policy (ISSP) [28]. The Security Policy describes the NETC@RDS information system security needs and requirements and provides the basis for a secure operational environment.

Each partner must respect the ISSP to be allowed to access / connect its portal to other portals. It is also foreseen in the ISSP that security audits must be conducted each year to verify the NETC@RDS ISSP compliance [28]. A suitable security audit procedure and tool has also been constructed and approved. The audit procedure has been agreed and implemented by all partners.

The NETC@RDS security infrastructure has been divided in three layers: Layer 1 (international/country-to-country level, between national NETC@RDS portal servers), Layer 2 (national level, portal-to-backoffice), and Layer 3 (national level, portal-to-Service Units).

Layer 1 security implementation is common in all participating countries. At this level all national portal servers intercommunicate through the internet using PKI infrastructure to achieve a high level of secure communication. All servers should have a reverse proxy service in place that redirects communication either to other national portals or to local BackOffice services. All communication should also be SSL v3 encrypted and performed under HTTPS protocol. Additionally server identification and authentication should also be achieved using certificates. In other words, portal-to-portal communication should involve server and client certificate exchange in order to attain server identification, authentication and authorization. Layer 2 security layer refers to the security infrastructure that is in place in the NETC@RDS infrastructure between the national portal server and the back office services, while layer 3 refers to the security infrastructure that is in place in the NETC@RDS infrastructure between the national portal server and the Service Unit workstations (Service Unit points).

The NETC@RDS common security policy has been constructed under the basic principle that the network build among the NETC@RDS partners should not add any unacceptable new risk within any partner organization. In addition, appropriate technologies and procedures must be used to ensure that data travels with adequate safety over the network build among the NETC@RDS partners and is only disclosed to authorized parties**.** The NETC@RDS information security policy should also provide means of proof and essential checks which give users trust in the given information. It should also help establish the basic security requirements that must be satisfied in order to ensure system continuity and prevent and minimise the impact of security incidents by implementing a stable, reliable and secure infrastructure. Finally, the NETC@RDS security policy is constructed under the principle of well-proportioned answer to the incurred risk.

Regarding its context, the NETC@RDS ISSP recognises three main actors that interact in the NETC@RDS system: the end users group, the health insurance organizations and the national access point providers. It also recognises two main beneficiaries (the insurees and the liaison offices), and four main data exchanges (Authentication Data, Administrative Data, Validation data and Data with Test Values) (figure 4).

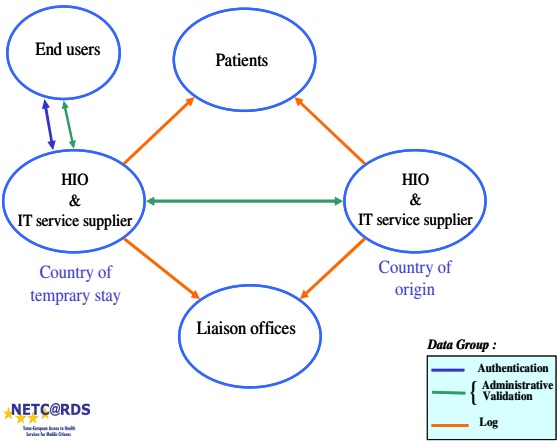


Fig. 4. Exchanges of data in the NETC@RDS context

Finally, as far as the legal basis is concerned, there are two distinguished dataflows identified in the ISSP for the NETC@RDS system: the NATIONAL dataflow level (from the End User to his National HIO/IT service supplier, or other related national dataflows) and the INTERSTATE dataflow level (from the National HIO/IT Service of the Member State of Temporary Stay, to the National HIO/IT service supplier of the Member State of Origin). For the NATIONAL dataflow, the actors should respect the respective national laws on data protection in effect, while for the INTERSTATE dataflow, as a pan European network, the NETC@RDS actors should respect at least the related European legislation (for example the European Directive 95/46/EC on data protection) [28].

The NETC@RDS ISSP includes 7 basic security rules (3 of national and 4 of European competency). It also includes a long list of procedural and technical recommendations.

A secure network interconnection between the national portals can also be provided by the integration in the EESSI architecture [1]. The respective national health insurance networks will be connected in this case by establishing national portals, which connect with each other via EESSI (figure 5). A cross-border electronic authentication request will then be routed through this network. EESSI is devoted to asynchronous cross-border data exchange between social security organizations, while NETC@RDS provides real time authentication mechanisms by on-line control between health practitioners and the foreign competent institution. However, this mechanism can be considered as a generic one and can be also adapted to other e-Gov/e-Health services like e.g. cross-border ID Management.

3.4 Example NETC@RDS Scenario

One of the most important purposes of NETC@RDS and the e-EHIC is to prove, through secure online cross-border authentication, the entitlement of a European

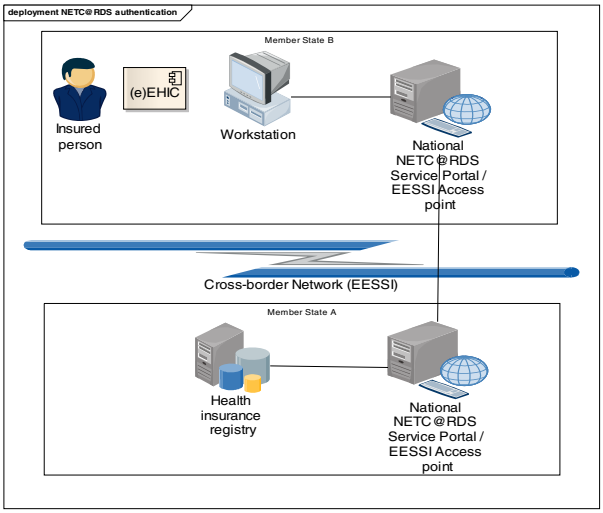


Fig. 5. Integration of NETC@RDS with the EESSI[2] architecture

citizen outside his/her home Member-State while requesting healthcare services. This can be described, via an example scenario, as follows (figure 6):

A French citizen is on vacation in Germany and needs to use unplanned healthcare services, i.e. the visitor goes to a German doctor because of sickness or maternity. For her entitlement, she shows either her eye-readable EHIC or, if the EHIC is expired, her electronic national/regional Health Insurance Card and provides it at the front desk at the doctor's facility. The card, containing the eEHIC dataset, is read by a smart card reader connected to the front desk workstation. This workstation connects to the national German NETC@RDS Service Portal via online connection and tries to verify the dataset. To this end it is necessary to authenticate the German doctor to this portal. The German NETC@RDS Service Portal then contacts the French NETC@RDS Service Portal, which in turn contacts the French Health Insurance company back office database for verification of the dataset and for authentication of the health insurance smart card shown as proof of entitlement at the point of health care delivery. This verification of entitlement contains the actual electronic authentication as a first step. The result of this verification is the decision (yes/no) about the entitlement of the patient, which is transmitted back to the front desk workstation of the German doctor.

It is possible to use either EHIC or different types of electronic national/regional health insurance cards. Independent of what type of card is presented at the patient check-in front desk in a hospital or in an ambulatory facility, the technical infrastructure enables first the capture of the EHIC dataset from various portable documents (i.e. a valid eye-readable EHIC or a national/regional health and insurance smart card or any other ID token) and then the validity of entitlement by the issuing institution.
SELECTED STUDIES

[2] EESSI = Electronic Exchange of Social Security Information [5,6].

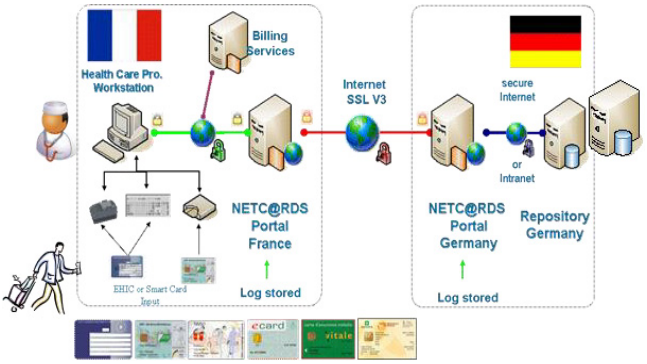


Fig. 6. How the existing NETC@RDS system works – example scenario

4 Mapping the ENISA Generic Model to the NETC@RDS Implementation

The NETC@RDS technical infrastructure using EESSI is mapped below to the proposed ENISA generic cross-border electronic authentication model, therefore proving that it can provide the basis for the implementation of the generic model. The NETC@RDS model including this mapping is described in figure 7.

Mapping the generic model to NETC@RDS is based on the following relations [1,2,28]:

The insured person residing outside his/her home Member-State requesting health services is the generic model's user. This user is entitled to obtain the services according to regulation 1408/71 [30]. The eEHIC maps to the token. The minimum set of data held on the token is prescribed in the Administrative Decision No 189 of 18th June 2003 [5]. The workstation within a hospital or ambulatory facility represents primarily the User Service Provider of the generic model. This workstation also reads the EHIC dataset from the token. This part of the workstation's hardware and software realizes the Token Adapter. The part of the NETC@RDS workstation that interfaces with the NETC@RDS Service Portal must be considered to be a first subcomponent of the Cross-Border Adapter [1].

While following this approach, the NETC@RDS project would follow the European Regulations for 883/04 on the coordination of social security systems [4]. Additionally to each Portal, national laws and regulations are applicable.

5 Addressing Cross-Border Protection Requirements with the Netc@rds Implementation Approach

The following case study provides an example of addressing NEC@RDS cross-border data protection requirements that demonstrates the suitability of NEC@RDS approach

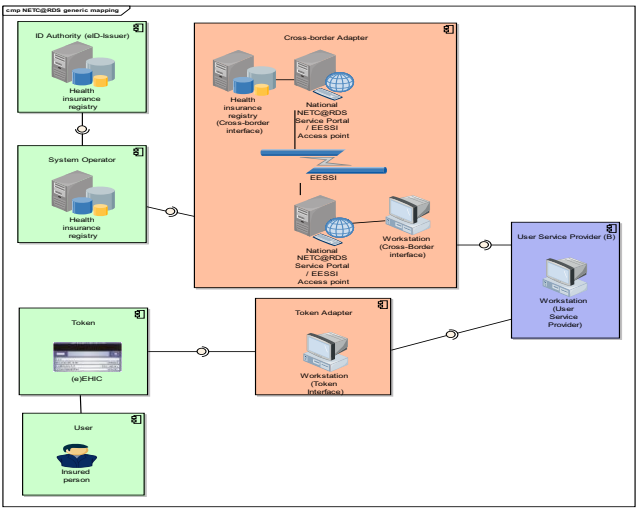


Fig. 7. Mapping of NETC@RDS components to the generic ENISA model [1]

for implementing the ENISA generic security model. It is based on the last evolution step of NETC@RDS in which the national service portals will communicate with each other securely via the EESSI [1]. In this function it is assumed that each national portal authenticates its domestic communication partners as persons or institutions authorized to request an electronic authentication. The national service portals and health insurance registers in other countries are not required to electronically authenticate the requesting health professional or institution as entitled to perform an authentication request, but can and must rely on the functioning of the first country's national service portal.

i. **Identity Data.** The Identity Data used and transmitted in the NETC@RDS electronic authentication is defined in the CEN Workshop Agreement CWA 15974 (May 2009) [25]. According to this document the EHIC identity data comprises the following mandatory information: surname of the card holder ("Name" on the face of the EHIC card), forename of the card holder ("Given names" on the face of the EHIC card), personal identification number of the card holder, date of birth of the card holder, expiry date of the card, ISO code of the Member-State issuing the card, identification number and acronym of the competent institution, logical number of the card (including a card issuer identifier), and identification of the paper form that is replaced by the card. The EHIC data set is transmitted to the Health Insurance Registry (the System Operator) during authentication.

ii. **Application Data.** In addition to the EHIC data set, other information is also transmitted during the authentication of an entitlement. This data comprises identification data on the health care professional and his institution (the User Service Provider), return codes and additional entitlement data.

iii. **eEHIC (Token).** The eEHIC contains the Personal Data. This data is defined to be freely readable. An authentication mechanism for the eEHIC may be

implemented optionally, but this must not hinder the free access to the eEHIC dataset. The mandatory EHIC dataset is also printed on the eEHIC surface. The eEHIC is under the control of the user, and it is assumed that the user consents to reading the data by handing the eEHIC to somebody.

iv. **Health Insurance Register (System Operator).** The Health Insurance Register or the IT systems of the health insurance company hosts the personal data for a large amount of users of the system. Thus large scale abuse of personal data is possible. The confidentiality of this data must be protected. The integrity of this data and of any additional application data must be ensured in order to allow the correct functioning of the system. Nevertheless these aspects are beyond the scope of a risk assessment for cross-border authentication, since the health insurance company is required to maintain the required levels of security also in its regular domestic and non-electronic cross-border operations. One main concern of the health insurance company as a stakeholder and participant in the NETC@RDS system must be that the introduction of this system must not compromise the company's established security levels.

v. **Workstation (User Service Provider).** The workstation at the medical institution has the primary function of allowing the medical institution to provide and account for services within the respective national health care system. This functionality must not be compromised by extending the workstation's tasks to accommodate the NETC@RDS system. The evaluation of security threats and protection requirements is limited to the functionality of the workstation that concerns the processing and storing of data related to the NETC@RDS system. The primary function of the workstation may pose other (higher) requirements.

vi. **Workstation (Cross-Border Adapter).** Software and potentially hardware must be added to the (domestic) workstation in the health care institution and the associated local IT systems in order to allow the cross-border authentication within the scope of the NETC@RDS system. These components are considered part of the Cross-Border Adapter and are governed by the respective local laws and contracts. It is assumed that the communication with the national service portal is performed via a secure connection that requires mutual authentication.

vii. **National Service Portal (Cross-Border Adapter).** The National Service Portal is the national focal point for all NETC@RDS cross-border activities. It is the interface between the national network and the European network EESSI. One main task of this portal is the authentication of health professionals and medical institutions to authorize the authentication request to the foreign health insurance registry. The National Service Portal passes authentication requests from domestic medical institutions across the border and receives authentication requests from abroad to be passed to the domestic Health insurance registers.

6 Limitations and Future Research Directions

As seen above, the proposed NETC@RDS approach can provide a suitable basis for a secure electronic cross-border authentication implementation, based on the ENISA generic security model, that suitably addresses the cross-border authentication requirements. There are still however a number of limitations and areas for further

research that need to be addressed if the proposed implementation approach is to be used as a multipurpose electronic cross-border authentication system.

Data privacy must be adequately protected in any such approach to electronic authentication, be it domestic or cross-border. Cross-border activity is governed however by the different laws and regulations of participating states. These laws often either affect or even prohibit specific transactions or data exchanges. Therefore, there is a need to further clarify the approach on how to respect the European and national data protection laws and regulations within the proposed electronic cross-border authentication system. There is, for instance, a need to further review the relevant EU and national law and regulation for impact on the design of the authentication system and also analyze additional legal provisions and regulations at national level. The regulatory basis for the cross-border implementation also needs to be further studied and incorporated. More detailed guidance should also be given to participating countries on how to best establish or extend existing mechanisms that meet the necessary level of trust.

The problem/risk of identity theft in electronic cross-border authentication also needs to be further studied. Further research is required to fully ensure that the eID token is used by its rightful holder and that the request for authentication is really in accordance to the will of a trustworthy authority / holder.

Cross-border authentication must also mutually establish, beyond any reasonable doubt, the identities of the user and the user service provider. To this end, a sufficient chain of trust must be established through all participants in the cross-border authentication process. The issue of how the system operator will establish sufficient trust in the identity of a user service provider across borders poses another interesting concern. The question mainly lies on the reliability and confidence with which a national portal authenticates its participants. The proposed common security policy for all participants in the cross-border exchange could provide the basis for such a suitable common level of security by all participants.

The cross-border system must also effectively ensure secure communications. These can either rely on secured publicly accessible internet connections or be integrated in dedicated secure cross-border networks (as for example EESSI). The authentication of the participants within the overall communication and the security of communication itself (e.g. by sufficiently strong encryption), also warrants further study.

Finally, the implementation of a secure real-time solution, based on existing EESSI flows, as for example the real-time verification of the patient's cross-border entitlement verification at the healthcare providers sites, needs to be further investigated, if the potential of EESSI is to be fully realized. This can be based on both the EHIC or national/regional social security cards and other portable ID documents.

7 Conclusions

Current electronically readable identity documents issued to citizens are usually available today only in a national context. European citizens moving freely through

Member States face the problem that their eIDs from their home state do not allow access to services of another Member State in which they are temporary present. It is therefore an important task to improve the cross-border interoperability of electronic identification and authentication systems, something that has been clearly highlighted in the ENISA report on secure cross-border automated services.

As seen in this paper, the NETC@RDS approach can provide a suitable, ENISA-model-based implementation approach for cross-border authentication that can also be applied in several application areas, including the social and health insurance sector. More specifically, it has been shown that the NETC@RDS approach can provide an implementation methodology for the ENISA generic model that can support electronic cross-border ID verification and authentication at an intra-European scale. It also helps address a number of other related risks related to electronic cross-border authentication, such as legal and regulatory issues, improvement of user credentials and bridging technological infrastructures at the national level, on which however further research is required.

References

1. Security Issues in Cross-border Electronic Authentication - ENISA (February 2010), http://www.enisa.europa.eu/activities/identity-and-trust/privacy-and-trust/eid/xborderauth
2. The NETC@RDS project, http://www.netcards.eu
3. European Union: Directive 2006/123/EC of the European Parliament and of the Council of 12 December 2006 on services in the internal market
4. European Union: Regulation (EC) No 883/2004 of the European Parliament and of the Council of 29 April 2004 on the coordination of social security systems
5. European Union: Administrative Commission of the European Communities on Social Security for Migrant Workers – Decision No 189 of 18 June 2003
6. European Union: Administrative Commission of the European Communities on Social Security for Migrant Workers – Decision No 190 of 18 June 2003
7. ENISA, 2008: Mapping IDABC Authentication Assurance Levels to SAML v2.0 – Gap analysis and recommendations
8. ENISA, 2009: Report on the state of pan-European eIDM initiatives
9. ENISA, 2009: Privacy Features of European eID Card Specifications
10. ICAO: DOC 9303 Part 1 Volume 1, Passports with Machine Readable Data Stored in Optical Character Recognition Format
11. ICAO: DOC 9303 Part 1 Volume 2, Specifications for Electronically Enabled Passports with Biometric Identification Capability
12. ICAO: PKD Regulations for the ICAO Public Key Directory
13. ICAO: Memorandum of Understanding regarding Participation and Cost Sharing in the electronic Machine Readable Travel Documents ICAO Public Key Directory (ICAO PKD)
14. ICAO: PKD Procedures for the ICAO Public Key Directory
15. ICAO: ICAO PKD Terms and Conditions
16. Hartmann, Körting, Käthler: A Primer on the ICAO Public Key Directory (2009)
17. Bundesamt für Sicherheit in der Informationstechnik: BSI Standard 100-1Infomation Security Management Systems (ISMS)

18. Bundesamt für Sicherheit in der Informationstechnik: BSI Standard 100-2 IT-Grundschutz Methodology
19. Bundesamt für Sicherheit in der Informationstechnik: BSI Standard 100-3 Risk Analysis based on IT-Grundschutz
20. IDABC: Common specifications for eID interoperability in the eGovernment context, http://ec.europa.eu/idabc/en/document/6484/5938
21. ICT-PSP STORK: D2.1 - Framework Mapping of Technical/Organisational Issues to a Quality Scheme, http://www.eid-stork.eu/index.php?option=com_processes&Itemid=60&act=streamDocument&did=579
22. ICT-PSP STORK: D2.2 – Report on Legal Interoperability, http://www.eid-stork.eu/index.php?option=com_processes&Itemid=60&act=streamDocument&did=578
23. ICT-PSP STORK: D2.3 - Quality authenticator scheme, http://www.eid-stork.eu/index.php?option=com_processes&Itemid=60&act=streamDocument&did=577
24. ICT-PSP STORK: D4.1 Interim Report on eID Process Flows, http://www.eid-stork.eu/index.php?option=com_processes&Itemid=60&act=streamDocument&did=576
25. CEN: CWA 15974:2009 (E) Interoperability of the electronic European Health Insurance Cards (WS/eEHIC)
26. Sušelj, M., Zuffada, R.: Netc@rds for e-EHIC - a Step Towards the Introduction of the European Health Insurance Card (2005)
27. ISO/IEC 27002 Information technology - Security techniques - Code of practice for information security management
28. The NETC@RDS Security Policy, Deliverable D.5 (2011), http://www.netcards.eu
29. IDABC: EESSI (Electronic Exchange of Social Security Information) Website, http://ec.europa.eu/idabc/en/document/7189/
30. European Union: Council Regulation (EC) No 1408/71 of 14 June 1971 on the application of social security schemes to employed persons, to self-employed persons and to members of theier families moving within the Community
31. European Community: Decision No. 189 of 18 June 2003 of the Administrative Commision of the European Communities on Social Security for Migrant Workers
32. HPRO Card: Website, http://hprocard.eu
33. STORK: Website, http://www.eid-stork.eu
34. ICAO: ICAO PKD Interface Specifications
35. Council of Europe: The European Convention on Human Rights and its five Protocols, Rome 4 November 1950
36. European Union: Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data
37. European Union: Directive 2005/36/EC of the European Parliament and of the Council of 7 September 2005 on the recognition of professional qualifications
38. EU Data Protection Directive, http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf
39. The ENED network, http://www.ened.eu

SWOT Analysis of a Portuguese Electronic Health Record

Rui Pereira[1], Maria Salazar[2], António Abelha[1], and José Machado[1]

[1] DI-CCTC, Universidade do Minho, Braga, Portugal
[2] Centro Hospitalar do Porto EPE, Oporto, Portugal

Abstract. In this paper it is describe a SWOT analysis of an Electronic Heath Record (EHR) implemented in a Portuguese hospital. As the EHR is a core part of a hospital information system, it is extremely important to ensure that it offers the best functionalities and that users are satisfied. With this analysis it is intended to gather information about the system, in order to improve the EHR implemented in the hospital. In the end, and appending to the results of a usability evaluation done in previous works, the evaluation team had enough knowledge about are the strengths and weaknesses of the EHR, as well as what opportunities can be taken and the threats that have to be avoided.

Keywords: Electronic Health Record, SWOT analysis and TOWS analysis.

1 Introduction

The sustained demand by healthcare organizations to improve the quality of patient care and patient safety boosted the adoption of information and communication technologies (ICT). Therefore, nowadays hundreds of ICT systems, such as the electronic health record (EHR), have been adopted in order to serve physicians as well as other professionals in their daily work with patients. An EHR can be defined as a longitudinal electronic record of patient heath information generated by one or more encounters in any care delivery setting, such as progress notes, medications, vital signs, past medical history, laboratory data or radiology reports. It has the ability to generate a complete record of a clinical patient encounter, which includes evidence-based decisions support, quality management and outcomes reporting. The EHR has proved to be an excellent tool for healthcare organizations. Once EHR systems play an important role in a hospital environment, it is vital to ensure that it presents the best conditions possible. In this context, this article highlights one manner to perform an evaluation to the state of this type of systems. The EHR implemented at Centro Hospitalar do Porto (CHP), in Portugal was subjected to a SWOT analysis in order to ascertain what can be change to improve the system. This analysis can reveal what are the great strengths of the system as well as its major pitfalls. In addition to this, the opportunities than can be taken advantage of are highlighted and the key threats to the system are alerted. At the end of this analysis, a great amount of information about the EHR is gather, which can be used to improve the system, availing to become a better tool for the professionals of the hospital as well as for patients.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 169–177, 2013.

2 The Electronic Health Record

The EHR, assumed as a Hospital Information System (HIS) for excellence, is a core application which covers horizontally the virtual health care unit and makes possible a transverse analysis of medical records along the several services, units or treated pathologies, bringing to healthcare units new computation models, technologies and tools, based on data warehouses, agents, multi-agent systems and ambient intelligence [1]. It receives the information of patients, in particular exams, thru a platform which main goal is providing rye interoperability between heterogeneous information system and medical equipment [8]. Beyond the organizational, functional, technical and scientific requisites, one may have to attend ethical and legal needs, as well as data quality, information security, access control and privacy. Despite the fact that there is not one exact definition for EHR, it can be defined as the computerized records of patients' clinical data. This information, which can be clinical, administrative or financial, is inserted in an electronic system that enables the capture, maintenance, transmission and storage of clinical information which is essential, not only for the monitoring of the health status of each patient but also for proposes such as cost management [2][3]. Thus, the EHR is an assembly of standardized documents, ordered and concise, directed to the register of information that can be compiled either by physicians or other health professional; a register of compiled facts, containing all the information regarding patient health data; and a follow up of the risks values and clinical profile [1]. The main goal is to replace hard documents by electronic ones, increasing data processing and reducing time and costs. The patient assistance will be more effective, faster and quality will be improved. With the adoption of such a system like EHR, it is possible to facilitate and improve care in health establishments. It enables the possibility of acquiring the versatility of a device capable of storing a vast sum of information, which can become more accentuated with the development of computer science. The data legibility and duplication, the continuous data processing, the ability to detect errors, the reduce frequency of loss records, the support communication between external sources of medical information, management and resource planning or releasing alarms concerning eventual pathological anomalies are some of the great advantages acquired with this system. Analysing the advantages at a structural level, the EHR supports the customization of the user interface, allowing the use of different layouts of insertion and viewing information under the very useful aspect of the availability of specific modalities in the hospital. Along with this, the EHR enables the automated collection of clinical parameters from monitors, imaging equipment, chemical analysis, among others [2]. Furthermore, it allows that the entire information can be share among different users whom are directly involved in the healthcare of the patient.

3 SWOT Analysis

The assignment of providing a well-functioning health service demands continual adjustments and sometimes also the introduction of new financing and organizational

methods. We must be careful, however, that in enterprise such modifications we do not obliterate aspects that are function well. Hence, it is essential to comprehend the assets and drawbacks of the EHR, as well as the prospects available to the system and the vulnerabilities that threaten it. In another words: Which features of the EHR should be safeguarded from any modification? Which characteristics can be improved? With these deliberations in mind, the idea of enterprise a so-called SWOT analysis of the EHR implemented at the CHP was developed. SWOT is an acronym that stands for: Strengths, Weaknesses, Opportunities and Threats. The origin of this technique still remains a little dubious, with the vast majority assigns the development of this strategic planning tool to Albert Humphrey, between the years 1960 and 1970. SWOT analysis (Table 1) aims to identify the strengths and weaknesses of the case of study (normally an organization) and, at the same time, the opportunities and threats presents in the environment. Strengths represent the internal power that an organization possesses to compete against its rivals [4]. Weaknesses represent aspects that negatively impact product and/or service value with regards to customers or competitive environment [5]. Opportunities are defined as a set of conditions suitable for achieving certain goals at the right time, and threats are any improper event or force in the external environment that causes harm to the organization's strategy [6]. Afterwards the identification of these factors, strategies are developed which may build on the strengths, eliminate the weaknesses, exploit the opportunities and counter the threats. The first two variables (strengths and weaknesses) can be identified by an internal appraisal and the remaining aspects (opportunities and threats) by an external assessment [7]. Hereupon, it is possible to say that SWOT analysis considers the organizational environment, which is composed by many elements that organizations deal with and form complex cause-and-effect type of relationships with. Environment can be divided into two categories: internal and external environment. The internal environment is one that can be controlled by the organization/institutional and therefore is directly sensitive to the strategies formulated. It deals with internal factors within an organization in various areas such as management, culture, finance, research and development, staff, operational efficiency and capacity, technical frameworks and organizational structure. The external environment is not in control of the organization. It acts homogeneously on the whole of organizations engaging in the same market and the same area. This way, opportunities and threats influence equally the entire organizations, whose probability of impact should be handled by each company separately. Hence, the external appraisal scans the entire factors that take place outside the organization's boundary such as political, economic, cultural, social, technological and competitive environment with a view to identifying opportunities and threats [7]. A variation of SWOT analysis is the TOWS analysis, projected by Heinz Weihrich. In the TOWS analysis the various factors are identified and these are then paired e.g. an opportunity is matched with a strength, with the intention of stimulating a new strategic initiative. With the TOWS analysis, the information gathered with the SWOT analysis is exploited in order to assemble different strategies (Table 2). The S-O strategy uses the strengths to capitalize on opportunities and the S-T strategy uses the

strengths to prevent the threats. The W-O strategy aims to improve the weaknesses through opportunities. The W-T strategy is the most defensive strategy of the TOWS analysis. All the strategies that can be created, are proposed to avoid threats and, at the same time, minimizing the weaknesses.

Table 1. The SWOT Matrix

	Internal	External
Positive	Strengths	Opportunities
Negative	Weaknesses	Threats

Table 2. The TOWS Matrix

	Opportunities	Threats
Strengths	S-O Strategy	S-T Strategy
Weaknesses	W-O Strategy	W-T Strategy

4 SWOT Analysis of the EHR

In order to gather information to perform the SWOT analysis, the EHR implemented at CHP was object of an intense study. This study undertook the perception of the various aspects present in the SWOT analysis. The results of this study were consolidated and subsequently the items of SWOT analysis emerged which are mentioned hereinafter.

4.1 Strengths of the HER

- Power management of change in the system;
- Ability to personalize objects like interface;
- High availability and support full-time;
- High accessibility;
- Security;
- Technologically modern system;
- Ease of maintenance;
- Ease of use (usability);
- Credibility of the management team;
- Immediate access to detailed clinical information;
- Reports customized to meet the needs required;
- High computing power;
- Interoperability;
- Ability to remotely access the system.

4.2 Weaknesses of the EHR

- System documentation nonexistent;
- Graphical interface somewhat confusing;

- Necessity of paper documentation in some services of the CHP;
- Insufficient education and training of health professionals;
- Computers are old and therefore slow.

4.3 Opportunities to the EHR

- Ability to integrate with other applications;
- Ability to provide information via Internet;
- Ability to expand and sustain new services;
- Increasing importance of digital files;
- Government incentives;
- Extinction of paper use in the CHP;
- Modernization and organizational development;
- Projection of more efficient and usable interfaces;
- Developing better and more effective security protocols;
- Increasing expectation of citizens to obtain answers of clinical services faster and, at the same time, reliable;
- Use of mobile devices to access the system;
- Use of new technologies in order to enrich the system.

4.4 Threats to the EHR

- High degree of competition from other systems;
- Expansion of software companies for the health market;
- Competition / market pressure;
- Competition for scarce talented IT resources;
- Economic-financial crisis and subsequent financial constraints;
- Readiness to recover from disasters;
- Cyber attacks (hackers);
- System is based on Internet Explorer.

5 TOWS Analysis of the EHR

Once the SWOT analysis is concluded, it is possible to combine the outcomes from this analysis, promoting the TOWS analysis. Through this analysis it is possible to scrutinize the combinations between the different aspects of the system.

5.1 Strengths to Maximize Opportunities

- The possibility of modifying the system and customizing objects allows the projection of a more efficient and usable interface;
- The safety that the system grants coupled with the technical support that exists at full-time enables the development of better and more effective security protocols;

- Once the platform is at the forefront in terms of technology, numerous opportunities are glimpsed. First, the possibility of implementing the system in other hospitals, both nationally and at international level. However, organization modernization and the possibility of acquiring technologies at attractive prices may reinforce this status of the system, making it a viable alternative to similar systems from reputable companies;
- The remarkable interoperability that exists in the system permits the integration of new applications and the expansion and support of new services;
- The fact that the system can be accessed remotely facilitates the availability of information on the Internet (e.g., patients access their own information comfortably at home) and healthcare professionals access the system via mobile devices;
- The excellent specifications on the handling of the system (immediate access to detailed information, high-capacity computing, among others) based on the growing expectation that citizens have in obtaining responses of clinical services faster and at the same time reliable.

5.2 Strengths to Minimize Opportunities

- All the strengths verified in the system devaluate the competition, whatever it may be. However, it is essential to continue the improvement of the system, because the competition is expected to do the same;
- The security of the system restricts the exposure to computerized attacks and, consequently, violation of patient privacy;
- Since the system is updated, as the technology is concerned, the huge competition for talented and lacking IT resources does not have great influence on the system.

5.3 Opportunities to Minimize Weaknesses

- The projection of more efficient and usable interfaces guides for improving the current system interface;
- Government incentives, organizational modernization and new technologies with attractive prices are conditions that provide the upgrade of computers in the CHP, which are already a little outdated.

5.4 Minimization of Weaknesses and Avoid Threats

- There is not any system documentation, but is compensated by the presence of full-time technicians in the CHP;
- Modification of the system interface to not be overtaken by competition.

6 Discussion

Through the SWOT analysis, it was possible to find out that the existing EHR in CHP is a system of high relevance, owning innumerable positive characteristics. Aspects

such as interoperability, good usability and high availability of the system, foster the EHR implemented in CHP within such systems. However, it also has some weaknesses, though outnumbered when compared with its strengths. With this analysis, those weaknesses were identified and can be, now, overcome.

The absence of some kind of system documentation is addressed by the presence of full-time technicians, who are continuously available to assist any healthcare professional that encounters some sort of obstacle in the handling of the system. The interface is a tricky question. As the Portuguese legislation forces the healthcare professional to document in detail the entire information relatively to the patient, which can spawn a huge amount of information. Therefore, when a professional access the clinical process of a patient, every part of the patient's information has to be displayed in the screen, which can make the reading of the process a bit confusing. This aspect can be partially resolved by using screens of larger dimensions (widescreens).

The computerization of the entire clinical process in the wholly services of the CHP is not an easy task. However, it is one of the main goals and steps are being taken to achieve that desired ambition. It is not an easy task and it may take some time, but all the efforts will be compensated.

The question of the old computers is complicated since technology is constantly progress and follow that improvement throughout the entire hospitals is an expensive task indeed. Added to this, the Portuguese financial situation does not benefit the constant modernization. It is a question that will take its time to be overcome.

This analysis showed that there are some opportunities that can be exploited. For instance, the importance that digital files have obtained in recent years coupled with organizational modernization can boost the existing propose to turn the hospital paperless. New services, features or even newly developed systems can be integrating into the EHR, increasing the relevance of the current system. These and others opportunities that came up cannot be overlooked, once is vital to improve the current system in order to be up with the competition.

It is important to take into account possible threats to the EHR. This analysis shows up a few threats that we need to be aware. The faced competition is one of the largest threats, if not the greatest, that this system has to face. There are a lot of EHR's solutions, some of them from reputable companies, like Siemens. However, not only existing solutions is important to take into account, as more and more software companies want to engage into the hospital market.

The economic and financial crisis is another key threat. High financial constraints and fear to hold large investments may constrain the bet in new IT resources, which are scarce. IT resources that proved to be talented will raise the dispute for them, which can be another threat.

Security is one issue that is widely considered as a main threat. However, we totally trust in the security of the EHR implemented in the CHP. Nevertheless, it is important to ensure the security and confidentiality of its information, avoid potential cyber attacks and have alternatives to disaster situations such as, for example, the situation of the system crashing. If this happens, the CHP cannot paralyze the activities, and therefore CHP should have alternatives to flank the situation.

If the SWOT analysis provided the identification of the strengths, weaknesses, opportunities and threats of the system, the TOWS analysis unveiled strategies to improve the EHR system. The strengths of the system can maximize the existing opportunities. For example, as the systems grants interoperability, new services can be integrated into the EHR. The strengths can also be used to minimize the threats. One example of this strategy is the following. The high security stated in the system enables the restricted exposure to computerized attacks. Once the system presents a great amount of strengths, a lot of opportunities can be harnessed and threats avoided, as stated earlier.

This analysis can built on other two strategies. The opportunities can minimize the weaknesses of the systems. For example, the interface of the system was one of the weaknesses and there was one opportunity identified related to that aspect, the projection of more efficient and usable interfaces. Taking advantage of this opportunity, the system can be improved, and the flaw is overcome. The last approach seeks the minimization of the weaknesses of the system as well as avoiding the threats.

At the end of those analyses, the CHP acquired extremely valued information regarding the EHR system. At this point, they know what are their best practices, what has to be enhanced, what prospects have to be exploited and the dangers to circumvent. In addition to this, significant strategies were developed and the CHP can use them to improve the existing EHR.

7 Conclusions

In this paper it is presented a possible strategic planning to an EHR system, appending to the results of a usability evaluation done in previous works [9]. Considering that strategic planning has not been done to the EHR implemented in the CHP so far, such kind of plan was essential for the hospital. The results show that the system has a lot of strong points as well as fewer weak ones. With the identification of the system's weaknesses, it is possible to circumvent them. This evaluation proved to be an excellent tool, which has provided useful information to improve the quality of the EHR, relevant to offer better conditions not only to the healthcare providers but also to the patients.

Acknowledgement. This work is financed with the support of the Portuguese Foundation for Science and Technology (FCT), with the grant SFRH/BD/70549/2010 and within project PEst-OE/EEI/UI0752/2011.

References

1. Machado, J., Alves, V., Abelha, A., Neves, J.: Ambient intelligence via multiagent systems in the medical arena. Engineering Intelligent Systems for Electrical Enginering and Communications 15(3), 151–157 (2007)
2. Machado, J., Abelha, A., Novais, P., Neves, J., Neves, J.: Quality of Service in Healthcare Units. In: European Simulation and Modelling Conference, European Technol Inst. Le Havre, France, European and Simulation Modeling Conference 2008, pp. 291–298 (2008)

3. Neves, J., Santos, M., Machado, J., Abelha, A., Allegro, F.S., Salazar, M.: Electronic health records and decision support local and global perspectives. World Scientific Enginerring Academy an Society, Transactions on Biology and Biomedicine 5(8) (2008)
4. Sharplin, A.: Strategic Management. McGraw-Hill, Book Co. (1986)
5. Stahl, M.J.: Total Quality Management in a Global Environment, Oxford (1995)
6. Rowe, A.J.: Strategic Management: A Methodological Approach, 4th edn. Addison-Wesley Publishing Co. Inc.
7. Dyson, R.: Strategic development and SWOT analysis at the University of Warwick. European Journal of Operational Research 152(3), 631–640 (1994)
8. Peixoto, H., Santos, M., Abelha, A., Machado, J.: Intelligence in Interoperability with AIDA. In: Chen, L., Felfernig, A., Liu, J., Raś, Z.W. (eds.) ISMIS 2012. LNCS, vol. 7661, pp. 264–273. Springer, Heidelberg (2012)
9. Pereira, R., Duarte, J., Salazar, M., Santos, M., Neves, J., Abelha, A., Machado, J.: Usabiltiy Evaluation of an Electronic Health Record. In: Proceedings of the IEEE-EMBS International Conference on Biomedical Engineering and Sciences, Langkawi, Malaysia (2012)

Discovering Authentication Credentials in Volatile Memory of Android Mobile Devices

Dimitris Apostolopoulos, Giannis Marinakis, Christoforos Ntantogian, and Christos Xenakis

Department of Digital Systems, University of Piraeus
Piraeus, Greece
{mte1102,mte1116,dadoyan,xenakis}@unipi.gr

Abstract. This paper investigates whether authentication credentials in the volatile memory of Android mobile devices can be discovered using freely available tools. The experiments that we carried out for each application included two different sets: In the first set, our goal was to check if we could recover our own submitted credentials from the memory dump of the mobile device. In the second set of experiments, the goal was to find patterns that can indicate where the credentials are located in a memory dump of an Android device. The results revealed that the majority of the Android applications are vulnerable to credentials discovery even in case of applications that their security is critical, such as web banking and password manager applications.

Keywords: Android, Android applications, mobile security, volatile memory acquisition, credentials discovery.

1 Introduction

We live in a transitional period from desktop computing to ubiquitous computing. Modern mobile devices, such as smartphones and tablets, offer the capability of any time - any place computing. More than enough processing power, at least 512MB RAM, multicolor readable screens, GPS, Wi-Fi and 3G/4G internet access are some of the characteristics of today's mobile devices, putting virtually all the power of a desktop computer in a pocket. These capabilities allow people to carry with them the full contents of the Web along their whole digital life.

The proliferation of mobile devices has led to the birth of mobile digital forensics, a branch of digital forensics that deals with the recovery of digital evidence or data from a mobile device under forensically sound conditions. Most of the mobile devices' forensics research has been focused on areas like the acquisition and analysis of the internal flash NAND memory, SD Cards, understanding the file system and scrutinizing application files for malware analysis. However, little research has been done in the acquisition and analysis of the volatile memory (also referred as RAM) of mobile devices. This is the main reason that our work focus explicitly on the volatile memory of mobile devices.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 178–185, 2013.

The authors in [12] have demonstrated that forensic investigators can discover critical information in the volatile memory of desktop computers like users' authentication credentials. Therefore, an interesting question that arises is whether forensic investigators can discover sensitive data in the volatile memory of mobile devices. However, an important difference between desktop computers and mobile devices is that unlike desktop computers, the applications processes in mobile devices are not properly closed and continue to run in the background. This happens because the applications in order to be closed, the user must manually force to stop them. Thus, the volatile memory of mobile devices may contain a wealth of information which can be used as evidence in a court.

The previous statement is also interesting from a security point of view. It is a fact that mobile devices can be easily stolen or misplaced, due to their compact size. The loss of a mobile device can lead to major privacy breach, since emails, social activities and pictures can be disclosed using readily available tools. Symantec's lost cell phone study [1], reveals that people who found a smartphone, violated the loser's privacy a whopping 89% of the time. On nearly half of those phones (43%), the finder attempted to access the owner's online banking application. Considering also the fact that 61% of internet users reuse passwords on multiple websites/services [3], sometimes the disclosure of one password is enough to compromise the privacy of all the user's applications. Thus, from a security point of view, it is motivating to examine if a malicious can discover authentication credentials in the volatile memory of mobile devices and breach the privacy of the device's owner.

Driven by the above observations, this paper investigates whether authentication credentials in the volatile memory of mobile devices can be discovered using freely available tools. We focus on mobile devices that operate with the Android Operating System (OS), since it is one of the most widely used mobile OS [2, 4]. The experiments that we carried out for each application included two different sets: In the first set, our goal was to check if we could recover our own submitted credentials from the memory dump of the mobile device. In the second set of experiments, the goal was to find patterns that can indicate where the credentials are located in a memory dump of an Android device. The results revealed that the majority of the Android applications are vulnerable to credentials discovery even in case of applications that their security is critical, such as web banking and password manager applications.

The rest of this paper is organized as follows. Section 2 provides background information and the related work. Section 3 presents the carried out experiments by analyzing the methodology we followed and elaborating on the obtained results. Finally, section 4 concludes the paper.

2 Background

2.1 Android Mobile Operating System

Android is a Linux-based OS designed primarily for touchscreen mobile devices, such as smartphones and tablet computers. Android uses native open source C libraries to perform OS tasks and uses Java as a language for developing Android applications.

To run applications, Android uses its own Virtual Machine called Dalvik [5]. The latter creates Dalvik executable files (.dex), which are byte codes from .class and .jar files. The compact .dex format is designed to be suitable for systems that are constrained in terms of memory and processor speed. Each Android application runs in a separate process within its own Dalvik instance, relinquishing all responsibility for memory and process management to the Android run time, which stops and kills processes as necessary to manage resources [6].

Android devices have different types of memory used for different purposes. They have a volatile physical memory (i.e., RAM) that loses gradually its data when power is switched off. On the other hand, the internal storage memory of Android devices, which is based on NAND flash technology, is a type of non-volatile storage that does not require power to retain data. Moreover, SD cards store the file system of Android OS named YAFFS2, application files and multimedia files.

In many cases, Android applications can display and process data which are never saved or cached in non-volatile memory storage. For example, in banking and financial applications, passwords are not stored and the user of the device must type and submit his/her authentication credentials each time the application is used. Therefore, the credentials are only transmitted inside the device but not stored. This type of data is defined as data in motion [7].

2.2 Related Work

Research projects on acquiring and examining the volatile memory of Android devices (or in general of mobile devices) are very limited. In [11], the authors focus on specific running processes, and use the *ptrace* functionality of the kernel to dump specific memory regions of a process. However, the analysis of the virtual memory captures is limited in the sense that they are used only to discover evidences related to communication based applications (i.e., incoming and outcoming messages of chat applications). Moreover, a project named *Volatilitux* [8] provides only limited analysis capabilities, including enumeration of running processes, memory maps and open files, and does not provide a method to acquire memory from Android devices.

The most recent and concrete work on this area is found in [13]. The authors present a methodology for acquiring complete memory captures from Android by cross compiling a kernel and loading it into a rooted Android device. They also released a practical tool named *DMD* (now named *Lime Forensics*) [9] to demonstrate the feasibility and forensically soundness of their methodology. However, the authors did not analyze further the memory dumps to discover sensitive data.

Moreover, in [14] a technique is described for dumping the memory of an Android application by sending to the application the command *kill*. Next, analyzing the memory dump the author succeeded to discover an encryption key that used to decrypt SQLite databases containing sensitive data such as contacts, configuration settings, keys, etc. However, this method can be applied only to Android versions until 2.1, because from Android 2.2 and above the command *kill* has been removed.

Finally, the authors in [12] have described a memory acquisition process for desktop computers in order to recover passwords from volatile memory, even when the

computer is switched off. They examined some popular web application such as Facebook, Skype, and Gmail. In 56% of their tests they were able to recover passwords from the memory of the computer 15 minutes (in one case even 60 minutes) after the user disconnected and closed the computer. Moreover, they concluded that the time which someone can recover data depends on the type of RAM, but also by external factors such as the temperature of the room where the computer resides. Contrary to [12], our work focuses on Android devices, where the memory acquisition process presents significant challenges compared to the desktop computer counterparts [13].

3 Experiments

3.1 Methodology

To dump the volatile memory of Android devices we used an open source tool named the Dalvik Debug Monitor Server (DDMS) [10], which is a GUI based debugging application that allows the examination of running processes. Although its primary goal is to help developers to identify bugs in Android applications, we have successfully managed to use DDMS as a tool to dump the volatile memory of a running process.

In a six months period we examined thirty Android applications. The majority of these applications release updates frequently. We mention that all experiments were performed in the latest version of the applications until October 1st, 2012. Based on their functionality, we divide the applications into four categories. The first category consists of m-banking applications, the second category includes e-shopping/financial applications, the third category includes social networks and communication applications, and finally, the fourth category consists of password manager and encryption applications. The applications of the first three categories simply use a username and password to authenticate a user. On the other hand, the applications that belong to the fourth category (i.e., password managers and encryption applications) use the password of the user (or the concatenation of the password with a random string to increase entropy) as a key to encrypt/decrypt passwords of other applications or any other sensitive data.

The experiments that we carried out for each Android application included two different sets: In the first set, our goal was to check if we could recover our own submitted credentials from the memory dump of the mobile device. In the second set of experiments, the goal was to find patterns that will indicate where the credentials are located in a memory dump of an Android device. This would be beneficial in forensics, where the investigator has a memory image of an Android device and he/she may use these patterns to find unknown credentials. On the other hand, as a negative side effect a malicious can stole an Android device and try to disclose the passwords of the owner of the device.

A problem that we had to address was the fact that the majority of mobile devices are shipped with a custom version of the Android OS designed from the mobile device manufacturers. Therefore, the memory dumps can be different for each Android device. To cope with this issue, we decided to experiment in an Android emulator that

uses pure versions of the Android OS. However, to verify our findings from the emulator we have also experimented in a few actual Android devices (i.e., Sony Ericsson Xperia x8, Samsung Galaxy S / S Plus, Motorola Xoom Tablet).

In the first set of experiments we performed the following steps: First, we installed the applications under investigation for the various versions of the Android OS. Next, we chose and open randomly an application and we submitted our own credentials (i.e., username and/or password) to log in. Then we set the application process to run in the background. This was done because most users instead of logging out and closing the application, they press the home button and return to the home screen. After this, we captured a snapshot of the volatile memory used by the specific application. To achieve this we used the DDMS tool that lists all the running processes. We selected the application process and we created a memory dump that includes all the data in motion used by the specific application. Finally, we searched for the users credentials, using the command-line tool *grep* combined with regular expressions to specify various possible combinations of the user's credentials.

The second set of experiments starts where the first set ends. In particular, after we located the credentials inside the memory dump, we tried to identify patterns which are nearby (in terms of bytes offset) to the identified credentials.

3.2 Results

Since we did not observe great discrepancies between the emulator and the actual devices, we do not present the results separately. In the first set of experiments, we successfully recovered our own submitted credentials in the majority of the applications, since they were in plaintext without any modification. The only modification that we observed in some applications was that the characters of the username and password strings had a dot symbol between them. For example, if the password of an application was submitted as "dssec", then we located the password as "d.s.s.e.c." (see Fig. 1). The reason of this trivial modification was due to Unicode encoding (i.e., UTF-16).

Regarding numerical results, in the first category (i.e., m-banking), we tested five mobile banking applications of five major Greek banks. We succeeded to retrieve the passwords in all five applications and the usernames in four of them. In the second category (i.e., e-shopping/financial), we tested three applications. We managed to retrieve the passwords in two applications and the usernames in all three applications. Regarding the third category (social networks and communications), we tested five applications. We successfully retrieved the username and password from all of them. Finally, in the fourth category (i.e., password managers and encryption) we tested seventeen applications. In this category, we managed to retrieve the passwords for all the applications. We summarize the results of our experiments in Table 1. In total, we succeeded to recover the passwords from twenty nine out of thirty examined applications and twelve usernames out thirteen applications that use usernames.

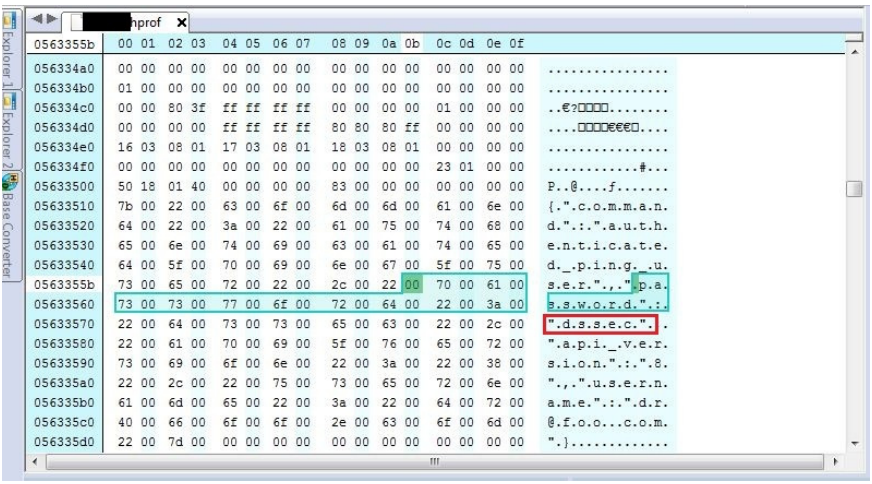


Fig. 1. Locating the password “dssec” and the pattern “password:” of an Android application in the memory dump file

In the second set of experiments, in all the applications we found a specific pattern that indicated the location of the password. For example, as shown in Fig. 1, right before the password that we had submitted (i.e., dssec), we found the string “password:”, evidently indicating that the following string will be a password. Some other patterns that we found were “username:”, “login name:” and “@Paw”. Therefore, a forensic investigator (or a malicious) can simply dump the volatile memory of an Android device and search for these patterns to discover the passwords of the owner of the device.

Table 1. Numerical results for each application category

	Usernames	Passwords
m-banking	4/5	5/5
e-shopping/financial	3/3	2/3
Social networks/communications	5/5	5/5
Password managers/encryption	-	17/17
Total	12/13	29/30

From the analysis of the previous results we can deduce that the majority of the Android applications are vulnerable to credentials discovery. It is alarming that even web banking applications proved to be vulnerable, leaving their costumers unprotected against password disclosure. Moreover, password managers that encrypt passwords of other applications were all vulnerable, since we found the master password in the volatile memory that is used to encrypt all other passwords. Thus, in case a user loses his/her Android device, his/her whole internet activity may be in jeopardy. Given the fact that many users use the same password for different applications, one password discovery is enough to breach the security of all the applications.

4 Conclusions

This paper investigated whether authentication credentials in the volatile memory of Android mobile devices can be discovered using freely available tools. This would be beneficial in forensics, where the investigator has a memory image of an Android device and he/she may use these patterns to find unknown credentials. On the other hand, as a negative side effect a malicious can stole an Android device and try to disclose the passwords of the owner of the device. The experiments that we carried out for each application included two different sets: In the first set, our goal was to check if we could recover our own submitted credentials from the memory dump of the mobile device. In the second set of experiments, the goal was to find patterns that will indicate where the credentials are located in a memory dump of an Android device. Results showed that revealed that the majority of the Android applications are vulnerable to credentials discovery. In particular, we succeeded to recover the passwords from twenty nine out of thirty examined applications and twelve usernames out 13 applications that use usernames.

As a future work, we will extend the experiments to cover scenarios where the goal is to discover authentication credentials in cases where the mobile device is set to sleep mode, when it is closed or even when the battery is removed. Moreover, in all these cases, we will investigate how much time the contents of the memory are preserved before they are completely faded out.

Acknowledgments. This work was sponsored in part by the project "Biotautotita" (GSRT 09SYN-72-597) of the framework "Synergasia", funded by the General Secretariat for Research and Technology Development Department (GSRT).

References

1. http://www.foxnews.com/tech/2012/03/12/symantecs-lost-cell-phone-study-confirms-worst-in-people/ (retrieved on November 2012)
2. https://www.google.com/nexus/4/#play (retrieved on November 2012)
3. Study of Consumer Password Habits (September 2012), http://www.csid.com/news/csid-conducts-study-of-consumer-password-habits-finds-disconnect-in-practices-and-mindset/ (retrieved on November 2012)
4. http://www.idc.com/getdoc.jsp?containerId=prUS23771812 (retrieved on November 2012)
5. Bornstein, D.: Dalvik VM Internals. In: Google I/O Developer Conference (June 2008)
6. http://mobworld.wordpress.com/2010/07/05/memory-management-in-Android/ (retrieved on November 2012)
7. Hoog, A.: Android Forensics: Investigation, Analysis, and Mobile Security for Google Android. Syngress, Elsevier (June 2011)
8. Girault, E.: Volatilitux: Physical memory analysis of Linux systems (December 2010)
9. http://code.google.com/p/lime-forensics/ (retrieved on November 2012)

10. http://developer.Android.com/tools/debugging/ddms.html (retrieved on November 2012)
11. Vrizlynn, T., Ng, K.Y., Chang, E.-C.: Live memory forensics of mobile phones. In: Digital Forensic Research Workshop (2010)
12. Karayianni, S., Katos, V., Georgiadis, C.K.: A framework for password harvesting from volatile memory. International Journal of Electronic Security and Digital Forensics 4(2-3), 154–163 (2012)
13. Sylvea, J., Caseb, A., Marzialeb, L., Richard, G.: Acquisition and analysis of volatile memory from Android devices. Digital Investigation 8(3-4), 175–184 (2012)
14. http://thomascannon.net/projects/android-reversing/ (retrieved on November 2012)

Conceptualizing Perceived Benefits and Investigating Its Role in Adoption of Tablet Computers among Newspaper Subscribers

Jari Salo[1], Sami Kajalo[2], Matti Mäntymäki[3], and A.K.M. Najmul Islam[3]

[1] Oulu Business School, Oulu University, Finland
Jari.Salo@oulu.fi
[2] Aalto School of Economics, Aalto University, Finland
Sami.Kajalo@aalto.fi
[3] Turku School of Economics, University of Turku, Finland
{Matti.Mantymaki,Najmul.Islam}@utu.fi

Abstract. This paper conceptualizes perceived benefit as a combination of eco-consciousness, social prestige, and extra services and proposes a research model for user acceptance of tablet computers. The model was tested among newspaper subscribers (n = 904) in one of the Nordic countries using Structural Equation Modeling (SEM).

The findings suggest that attitude toward tablet devices is predicted by eco-consciousness, ease of use, social prestige, and extra services. In turn, attitude along with intention to use predict intention to buy. Furthermore, ease of use was also found to predict extra services.

Keywords: Attitude, eco-consciousness, intention to buy, perceived benefit, tablet computers, technology acceptance model.

1 Introduction

Due to the advent of smartphones and tablet devices, the publishing industry has been in turmoil. To respond to the challenge, new services and new business models have emerged. Altogether, this has changed the traditional industry logic [9, 8]. Services are offered to different types of mobile devices e.g. tablet computers using digital technologies. For the publishing industry, digital technologies and mobile services provide both opportunities and challenges to excel in their business [40].

Since the 1990s newspaper publishers have felt pressured to have a web presence to be competitive on the market [9, 18]. Most players in the newspaper industry are either providing a replicate of its paper based newspaper or a modified version online [25]. One of the main reasons for the popularity of online newspapers is the possibility to utilize different technologies such as audio, graphics, video and interactive elements that are not possible to implement in a print version. For readers, on the other hand, interactive elements mean freedom to filter news according to their own preferences and the opportunity to give feedback on the articles, which creates a feeling of empowerment [37].

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 186–199, 2013.

The digital publishing market is expanding rapidly and new players such as the Daily have entered the newspaper market. Moreover, buying content online is increasing among consumers altogether fundamentally changing the way people consume newspapers. To understand the phenomena related to the diffusion (see Rogers, 2005 [34]) of tablet devices, it is essential to what drives the user acceptance.

As users of tablet devices are consuming increasing amount e-books and other mobile digital services, it is important to understand of tablet devices in order to accelerate their diffusion in the market [34]. Tablet devices are capable to offer several material and social benefits such as elevated social status, eco-friendly reading experience, and rich online content [8, 25]. Consequently, this research attempts to address the following research question:

How perceived benefit does influence the decision to adopt a tablet computer for e-reading?

User acceptance of tablet devices is an interesting research area because of the changing newspaper consumption patterns that urges the publishing industry publishers to identify new business opportunities.

In this study, we investigate the user acceptance of tablet devices using the Technology Acceptance Model (TAM) [11, 12] as a theoretical foundation for how different aspects of technology acceptance and features of tablet devices impact consumers' willingness to adopt these devices.

The paper proceeds as follows. In section 2 we present the theoretical background and develop the research model. Section 3 is dedicated to the data collection and data analysis methods. Section 4 presents a discussion on the data analysis findings and concludes with the implications and limitations as well as suggestion for further inquiry.

2 Theoretical Background and Research Model

2.1 Technology Acceptance Model

The Technology Acceptance Model [11] has been the most influential and most widely applied model for explaining users' intention regarding IT use [21]. TAM was originally developed from Ajzen & Fisbein's (1980) [1] theory of reasoned action (TRA).

TRA argues that if a person intends to commit a behavior it is likely that the person will actually do it. According to TRA, a person's behavioral intention depends on a person's attitude towards the behavior as well as the subjective norm [1]. Attitude is the combination of the user's beliefs about the consequences of performing the behavior and his or her evaluation of these consequences. The subjective norm is the combination of the relevant referents' expectations and the user's motivation to comply with these expectations.

TAM posits that perceived usefulness and perceived ease of use are the factors that determine an IS user's attitude, which, along with perceived usefulness, determines

behavioral intention. Finally, behavioral intention determines actual IS use. Perceived usefulness is described as being directly influenced by perceived ease of use.

TAM has been validated across time, population, and contexts [43]. TAM is well-established in Information Systems research as well as other domains [21]. Venkatesh et al. (2007) [43] argues that TAM has become nearly a law-like model and that it often serves as a basis for studies in other areas. For example, TAM and its constructs have been used in areas outside technology adoption such as information adoption [39], marketing [10], and advertising [33].

Several studies have focused on acceptance of mobile devices, mobile services, mobile advertising, and tablet computer adoption using TAM as the theoretical lens [8, 27, 35]. Majority of the studies found that benefit captured through perceived usefulness as the most dominating factor to predict adoption [8, 9].

Despite TAM's popularity, it has a number of limitations as discussed by leading IS researchers [4, 7, 17]. Specifically, there are two limitations of prior TAM based studies that motivated us to conduct this study. First, perceived usefulness is regarded as black box with little research effort into investigating different dimensions of usefulness [4]. Hence, TAM studies often lack actionable guidance for practitioners. Second, TAM assumes that more use is better and consequently TAM based studies focused on intention to use. In other words, more utilization of a technology increases purchasing. However, only a few prior adoption studies verify a relationship between use and purchase [26].

2.2 Conceptualizing Perceived Benefit

Prior technology adoption research conceptualized perceived benefit by perceived usefulness [11, 26, 42]. Perceived usefulness is an indicator of the degree to which the use of an IS will enhance a user's job performance [11]. Most of the prior studies ignored the uniqueness of research context and modeled perceived usefulness as a black-box focusing on an overall benefit of an IS [8, 44].

Many of the contemporary tablet devices can be used for other purposes than reading, including surfing the web, playing games and using social media applications which are here labeled as extra services. As argued in prior studies [6], the usage context of the IT artifact and the users [23] are important factors in explaining the adoption behaviors. As a result, in the present study we specifically focus on opening the perceived benefit and investigate its role in predicting the acceptance of tablet devices.

In the present study, we employ user's eco-consciousness, the social prestige, and extra services offered by the tablet devices in order to conceptualize perceived benefit. We employ these three benefits to conceptualize perceived benefit as these were argued as important in prior research [25]. Eco-consciousness is defined as the extent to which the consumer perceives using a tablet device as an ecological choice. Social prestige is defined as the degree to which use of the tablet devices is perceived to enhance one's status in his/her social system. Finally the construct, extra services is defined as the degree to which the tablet devices provide facilities to acquire additional information in a particular topic of interest.

2.3 Hypotheses Development

Figure 1 presents the hypothesized research model.

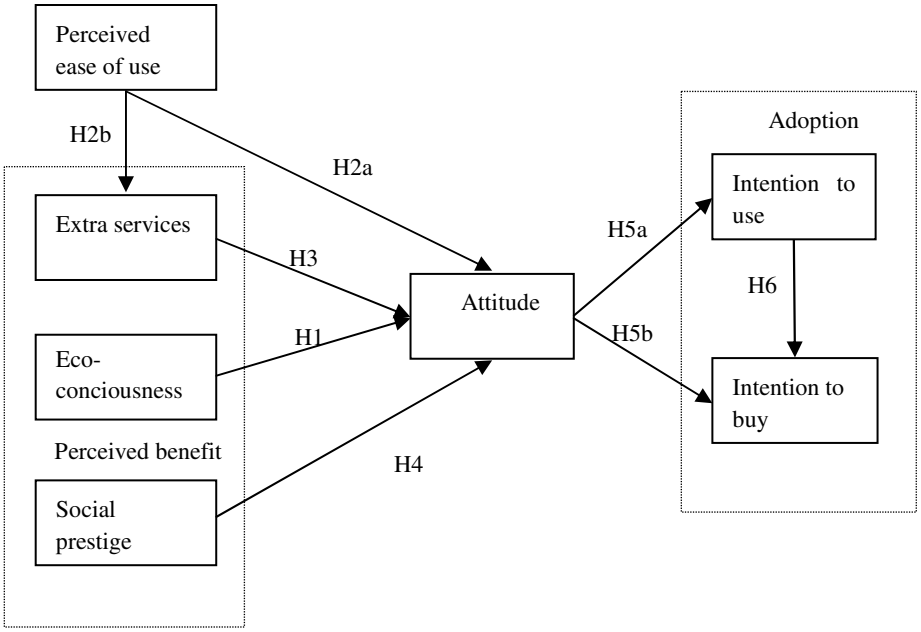


Fig. 1. Research model

Various environmental concerns are becoming increasingly prevalent and people tend to have more positive attitudes towards products that they perceive environmentally friendly [45]. Moberg et al. (2010) [25] found that using a tablet computer to read magazines and newspaper instead of having printed copies, the users may contribute to the environmental improvement. Following this finding, we expect that the users who perceive tablet devices as eco-friendly would have more positive attitude toward adopting a tablet device. In fact, eco-conscious users would avoid printed materials and prefer reading from the tablet screen. Prior research also argued that eco-friendly products can be considered status symbols and hence consumers are expected to have more positive attitude toward tablet devices [29]. As a result, we postulate the following hypotheses:

H1: Eco-consciousness is positively related to attitude towards using a tablet device.

According to TAM, we have positioned perceived ease of use as a predictor of attitude [11]. Perceived ease of use is defined as the degree to which a person believes that using a particular system would be free from effort [11]. Furthermore, it is plausible to assume that perceived ease of use determines how well the users can utilize the extra services of tablet devices [28]. Research on technology acceptance found that the extended use of an IS depends on the ease of use [19, 20]. Following

this, it can be argued that if a tablet device is easy to use, the users will use it for a wide variety of purposes. Conversely, if the tablet devices offer a lot of additional functionalities with a difficult interface, the users would hardly try to explore the additional functionalities. Instead they would prefer to use only the basic features. Thus, the hypotheses related to the role of perceived ease of use are postulated as follows:

H2a: Ease of use is positively related the attitude towards tablet devices.
H2b: Ease of use is positively related to extra services that tablet devices provide.

In reading books, newspapers and magazines, table devices offer several value-added features compared to the paper format. For example, articles can include links to complementary material e.g. audio, videos. Additionally, compared to traditional books and newspapers, tablet devices enable storing large amounts of material in a portable format so that the users can view the content at a later stage if they wish. Furthermore, the interesting news can be shared in the social media. Thus, we assert that these features of the tablet devices offer a major utility for the users and consequently positively affect the attitude towards using the device.

H3: Extra services are positively related to attitudes towards using a tablet device.

Attitudes do not develop in a vacuum but are influenced by e.g. other people's opinions, advertising and mass media [15]. As the diffusion of tablet computers has not proceeded into the late adopters and laggards stage, the use tablet computers can encompass certain status value for some users [34]. Prior technology adoption research has found image, which is similar to social prestige as a major driver of technology adoption [24, 42]. Thus we form the following hypothesis:

H4: Social prestige is positively related to the attitudes towards using a tablet device.

We argue that the actual adoption of a tablet device consists of two behaviors, usage and purchasing. Most prior research has investigated usage as the sole construct that reflect adoption [21, 44]. We argue that usage provides only a limited view of adoption. For example, many consumers may have the opportunity to use a tablet device owned by their family member of employer. They may or may not intent to buy a tablet device for his personal use. Thus, it is necessary to investigate both usage and purchase intention to understand tablet devices adoption. With this aim, we employed both use intention and buying intention in our research model.

Given that using a tablet device generally requires purchasing one, usage and purchasing can be viewed closely related behaviors with regard to the utilization of tablet devices. Grounded on TRA, TAM views behavioral intention being determined by attitudes [11]. As a result, we assert that attitude towards using a tablet device predicts both the use intention and purchasing intention.

H5a: Attitude towards using a tablet device is positively related to the intention to use a tablet device.
H5b: Attitude towards a tablet device is positively related to intention to buy a tablet device.

Finally, we argue that the two behavioral intentions, using a tablet device and purchasing one are related so that the use intention precedes purchasing. The argumentation is grounded on the assumption that process wise the willingness to use a shared (e.g., employer or family) tablet device would influence the users to buy their own devices at a later stage in order to fulfill a set of needs that takes place by using the device. Thus, we hypothesize the following:

H6: Intention to use a tablet device is positively related to intention to buy a tablet device.

3 Study Design and Method

3.1 Data Collection

The questionnaire was developed in several phases. First, the items of the questionnaire were developed on the basis of the literature review. Second, these questions were further developed after conducting several focus groups interviews among university students. Third, the item pool was pre-tested by sending the web survey to 200 newspaper subscribers. 23.5% of them answered the questionnaire and gave feedback that was used in developing the final questionnaire.

The data was collected through an internet survey in October 2010 among subscribers of a daily newspaper from one of the Nordic countries. It is a nationwide daily newspaper with long tradition and high standard of journalism. The link to the final questionnaire was sent by e-mail to 5000 subscribers of the newspaper. After one remainder e-mail 1084 out of 5000 respondents had answered the survey. Given the length of the web survey we find the response rate of 21.7% satisfactory. After filtering the incomplete responses, we ended up with 904 usable responses.

We consider the respondents represented well the subscribers of the newspaper. There were 454 male respondents and 450 female respondents, and thus there was a balance in terms of gender Moreover, no significant differences in means were found between early and late respondents on the scales studied (t-tests at .05 level), indicating that a non-response bias is unlikely to be a problem [3].

All measures in this study were collected via a single questionnaire, which opens up the threat of common method bias. Common method variance can either inflate or deflate observed relationships between constructs [31, 32]. We used the Harman's one-factor test to address this issue [32]. If a substantial amount of common method variance is present, either a single factor will emerge from the factor analysis, or one dominant factor will account for the majority of the covariance among the variables. All the items were entered into an exploratory factor analysis, using principal component analysis with varimax rotation. No dominant factor emerged which indicates the common method bias is not present in this study. The list of final measurement items in each scale of the inquiry is presented in Appendix 1.

3.2 Data Analysis

Structural equation modeling (SEM) technique has been used to test the hypothesized research model. SEM is particularly suitable for testing the proposed theoretical model, because it allows simultaneous estimation of multiple relationships between observed and latent constructs, and account for measurement error. Thus, we are able to test an overall model rather than just coefficients individually [5].

We used a two-step SEM approach in accordance with Anderson & Gerbing (1988) [2]. First, a confirmatory factor analysis (CFA) assessing the validity of the measurement model and the discriminant validity of the individual constructs was conducted. Second, a structural model is used to estimate the path coefficients and test for the relationships between constructs.

Following Anderson & Gerbing (1988) [2], CFA was carried out to assess the validity of all the construct measures included in this study. This assessment was carried out using LISREL 8.80 program. The fit of the measurement model was evaluated based on several goodness-of-fit indicators. These were the chi-square statistic ($\chi 2$), root mean square error of approximation (RMSEA), goodness of fit index (GFI), non-normed fit index (NNFI), and comparative fit index (CFI). RMSEA is usually regarded as the most informative of the fit indices, and values less than .08 are indicative of reasonable fit [36]. GFI is an absolute fit index, whereas NNFI and CFI are relative indices. For all three, values above .90 indicate good fit [36, 38]. The model under consideration provided good fit ($\chi 2$=813.11; df=188; p=.000; RMSEA=.055; GFI=.94; NNFI=.98; CFI=.98).

In order to evaluate the reliability of the latent variables, composite reliability values and average variance extracted values for each latent variable were calculated [13]. These are reported in Table 1 together with means and standard deviations for the scales, and the correlation matrix. They all exceeded the recommended level of 0.70 (ranging from 0.84 through 0.98). Likewise, average variance extracted values exceeded the recommended level of 0.50 (ranging from 0.57 through 0.92). Also, all factor loadings were statistically significant at the 5% level, and all of the factor loadings exceed the 0.50 level (ranging from 0.65 through 0.98) [16]. To conclude, our measures demonstrate adequate convergent validity.

Table 1. The scale means, standard deviations, reliability indexes and correlation matrix

Construct	Mean	S.D.	ρ_c	ρ_v	1	2	3	4	5	6
Eco-consciousness	3.58	1.72	.98	.92	1.00					
Ease of use	5.00	1.40	.92	.76	.11*	1.00				
Extra services	3.47	1.04	.84	.57	.22*	.28*	1.00			
Social prestige	1.59	0.94	.91	.78	.15*	-.04	.11*	1.00		
Attitude	3.94	1.66	.91	.77	.25*	.35*	.50*	.23*	1.00	
Intention to use	4.45	1.83	.94	.73	.22*	.38*	.51*	.12*	.77*	1.00
Intention to buy	2.76	1.64	.84	.67	.19*	.30*	.38*	.23*	.65*	.65*

S.D.: standard deviation, ρ_c: composite reliability, ρ_v: average variance extracted. ρ_c: $(\sum\lambda)^2/((\sum\lambda)^2+\sum(\theta))$,
ρ_v: $(\sum\lambda)^2/(\sum\lambda^2+\sum(\theta))$, λ: indicator loading, θ: indicator error variance. *=correlation is significant at .01 level (2-tailed).

The hypothesized model testing was undertaken by deploying covariance matrix and the maximum likelihood estimation procedure. Overall, the fit indexes (χ^2=959.27; df=199; p=.000; RMSEA=.059; GFI=.93; NNFI=.97; CFI=.98) imply a good model fit. Figure 2 illustrates the final structural model with standardized path estimates and t-values.

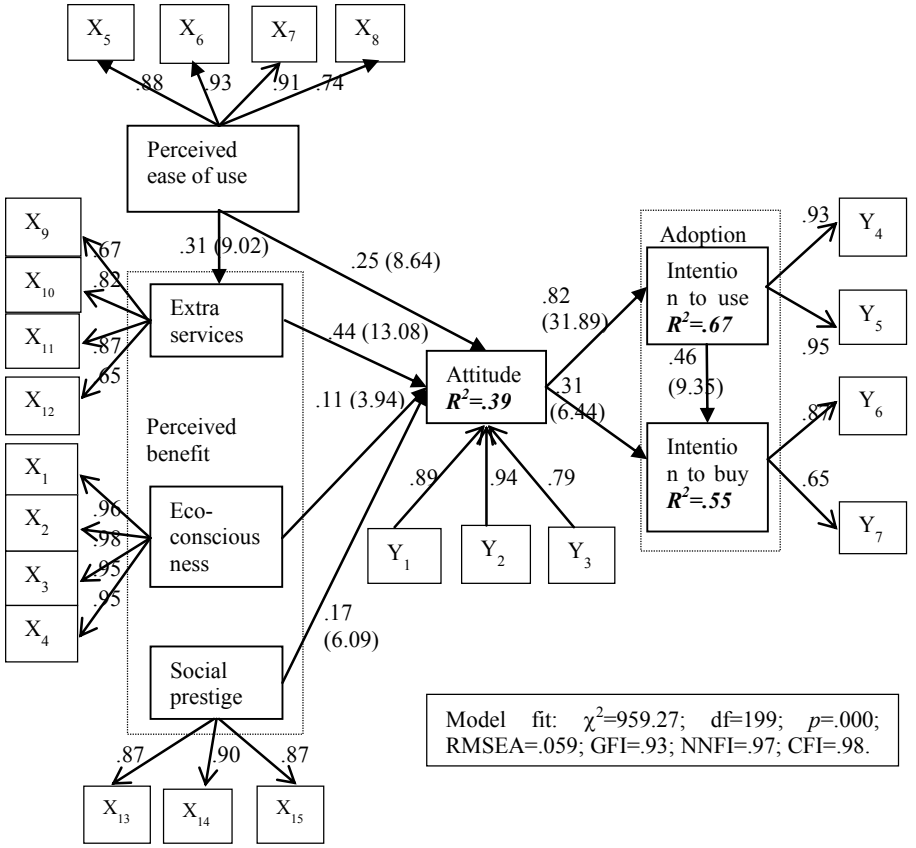


Fig. 2. SEM results

Figure 2 shows that Eco-consciousness is positively related to attitude (β=.11) created by tablet devices. These results provide empirical evidence for Hypothesis H1. However, these relations are rather weak.

Perceived ease of use is positively related both to attitudes towards tablet devices (β=.25) and extra services (β=.31) that tablet devices provide. These relationships are strong, and thus, strong empirical evidence for our hypotheses H2a and H2b is provided.

Also hypothesis H3 was supported, as extra services provided by tablet devices had a very strong positive influence on the attitudes towards tablet devices (β=.44). A weak relationship exists between social prestige created by tablet devices and the

attitudes towards tablet devices: the path coefficient was found at β=.17, supporting our hypothesis H4. Finally, the attitudes towards tablet devices is positively related to intention to use (β=.82) tablet device, and the intention to buy (β=.31) a tablet device. Also, as expected there was a strong direct relationship from intention to use to intention to buy (β=.46) a tablet device. These findings provide support for the hypotheses H5a, H5b and H6 respectively. To conclude, all our hypotheses received empirical support.

The explanatory power of the path model for each depended construct was examined by using R^2 values. These were 10 percent for extra services, 39 percent for the attitudes towards tablet devices, 67 percent for intention to use and, finally, 55 percent for intention to buy a tablet device.

To summarize, we found empirical evidence for all the hypotheses. The results suggest that ease of use and extra services that tablet computers provide are key elements when consumers develop attitude toward these devices. While ecological values and social prestige were found to have statistically significant impact on attitude, they were not as important as the ease of use and extra services that these devices provide.

4 Discussions

4.1 Main Findings

All of the proposed hypotheses were empirically supported. First, we found that all three constructs: eco-consciousness, social prestige, and extra services have significance influence on attitude. The relationship between eco-consciousness and attitude implies that environmental conscious users are likely to justify their immediate action based on their environmental knowledge [30]. However, it is interesting to observe that eco-consciousness was found to be a weak contributor in the adoption of tablet devices. This results suggest that users do not possess enough knowledge how using tablet computers for reading brings the environmental benefits.

The relationship between social prestige and attitude implies that tablet devices are regarded as status symbols. Prior technology adoption research used image as a variable focusing one's status symbol [24, 42]. Venkatesh & Davis (2000) [42] argued that the elevated social status provides a general basis for greater productivity. Following this line of argument, it is logical that social status influences the attitude toward using tablet devices. Thus, our finding is in line with prior literature.

It is interesting to observe that extra services played the strongest role in shaping users' attitude toward using tablet devices. This finding is supported by the media richness theory [14] which suggests a particular communication medium is preferred over another depending on the capability of the medium and users' task requirement. The online versions of newspapers definitely provide content in much richer format (e.g., videos, photos, hyperlinks) than that the printed versions. Thus, consumers consider the extra services as most important factor in order to develop their attitude toward the adoption of tablet computers.

Second, we found that perceived ease of use has significant influence on both extra services and attitude. The relationship between perceived ease of use and attitude is in line with the TAM tradition [11]. The relationship between perceived ease of use and extra services suggest that consumers choose to use the extra services only when the tablets are easy to use. This finding was not tested empirically in prior literature. However, some evidences can be drawn from the prior literature in support to this finding. For example, Hossain (2012) [19] found that many features of a complex information system often remain underutilized. Thus, it is expected that perceived ease of use would lead consumers use the extra services offered by the tablet devices.

Finally, we found that intention to use is predicted by attitude. This relationship is in line with the TAM tradition [11]. Furthermore, we found that both attitude and intention to use have significant influence on intention to buy a tablet device. In fact, it is interesting to observe that intention to use has the strongest influence on intention to buy. This result implies that consumers are more likely to purchase a tablet device if they would get a chance to use it beforehand.

4.2 Implications

Our research study has the following two major theoretical contributions. First, the study addresses how different aspects of technology acceptance and features of tablet devices impact consumers' willingness to adopt tablet devices. Given that tablet computers and dedicated e-reading devices are becoming increasingly popular, the present study makes a contribution to the research on the adoption of these technologies.

Second, we have conceptualized perceived benefit into three parts: eco-consciousness, social prestige, and extra services in the tablet computer adoption context. We have empirically tested the relationships of these variables with attitude toward using tablet devices. While most prior adoption research conceptualized perceived benefit as an overall benefit of using a system (i.e., perceived usefulness), our conceptualization and subsequent analysis sheds light on which benefit actually motivates consumers to adopt tablet devices for e-reading.

Our study findings have practical implications for tablet manufacturers and newspaper publishers. First, we found that eco-consciousness played a significant role for tablet adoption. Thus, the manufacturer needs to bring up the green values of using tablet computers for reading and avoiding printed media. As we found the influence of eco-consciousness on attitude as weak, we suggest manufacturer on describing how the tablet devices are eco-friendly in their marketing campaign.

Second, for newspaper industry, we recommend that only those devices that are easy-to-use are utilized as the main channel for content distribution. It might be that general tablet devices like iPad and Samsung Galaxy are most prominent for newspapers which have also color full content and can utilize embedded videos while single purpose devices such as Amazon Kindle could be more suitable for just e-reading.

For tablet device manufacturers it seems that easy-of-use is still the most important feature and to be aimed for. Furthermore, the role of tablet devices as ways for

reducing utilization of paper could be used as value based selling arguments in ads and marketing communication.

4.3 Limitations and Future Research

Our research has two major limitations that however also offer avenues for future research. First, the research was cross-sectional. The beliefs of the users regarding a system will change as the users gain experience of a target system but such changes cannot be captured with the type of cross-sectional study undertaken.

Second, our empirical analysis did not consider possible cultural differences, and different types of newspapers and magazines (e.g. local versus. national newspapers). Furthermore, we did not distinguish between different devices or operating systems (e.g. iPads vs. Android). Hence, we encourage empirical studies taking a comparative approach.

Third and finally, we did not examine the influence of price. Given that price is a key factor in setting the level of demand [44], future research incorporating the pricing of the newspaper and the potential value-added features would be managerially highly relevant.

References

1. Ajzen, J., Fishbein, M.: Understanding attitudes and predicting social behavior. Prentice-Hall, Englewood Cliffs (1980)
2. Anderson, J.C., Gerbing, D.W.: Structural equation modeling in practice - a review and recommended 2-step approach. Psychological Bulletin 103(3), 411–423 (1988)
3. Armstrong, J.S., Overton, T.S.: Estimating non-response bias in mail surveys. Journal of Marketing Research 14(3), 396–402 (1977)
4. Bagozzi, R.R.: The legacy of the technology acceptance model and a proposal for a paradigm shift. Journal of the Association of Information Systems 8(4), 244–254 (2007)
5. Bagozzi, R.P., Yi, Y.: Specification, evaluation, and interpretation of structural equation models. Journal of the Academy of Marketing Science 40(1), 8–34 (2012)
6. Benbasat, I., Zmud, R.W.: The identity crisis within the IS discipline: Defining and communicating the discipline's core properties. MIS Quarterly 27(2), 183–194 (2003)
7. Benbasat, I., Barki, H.: Quo vadis, TAM? Journal of the Association of Information Systems 8(4), 211–218 (2007)
8. Chen, S., Granitz, N.: Adoption, rejection, or convergence: Consumer attitudes toward book digitization. Journal of Business Research 65(8), 1219–1225 (2012)
9. Colleen, P.K., Chiagouris, L., Gopalkrishna, P.: Some people just want to read: The roles of age, interactivity, and perceived usefulness of print in the consumption of digital information products. Journal of Retailing and Consumer Services 19, 168–178 (2012)
10. Dabholkar, P.A., Bagozzi, R.P.: An attitudinal model of technology-based self-service: Moderating effects of consumer traits and situational factors. Journal of Academy of Marketing Science 30(3), 184–201 (2002)
11. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13(3), 319–340 (1989)

12. Davis, F.D., Bagozzi, R.P., Warshaw, P.R.: User acceptance of computer technology: A comparison of two theoretical models. Management Science 35(2), 982–1003 (1989)
13. Diamantopoulos, A., Siguaw, J.A.: Introducing LISREL: A guide for the uninitiated. Sage Publications, London (2000)
14. Daft, R.L., Lengel, R.H.: Organizational information requirements, media richness and structural design. Management Science 32(5), 554–571 (1986)
15. Entman, R.: How the media affect what people think: An information processing approach. The Journal of Politics 51(2), 347–370 (1989)
16. Fornell, C., Larcker, D.F.: Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 18(1), 39–50 (1981)
17. Goodhue, D.L.: Comment on Benbasat and Barki's "Quo vadis TAM" article. Journal of the Association of Information Systems 8(4), 219–222 (2007)
18. Greer, J.D., Mensing, D.: The evolution of online newspapers: A longitudinal content analysis, 1997-2003. In: Li, X. (ed.) Internet Newspapers: The Making of a Mainstream Medium, pp. 13–32. Lawrence Erlbaum Associates, New Jersey (2006)
19. Hossain, M.A.: Adoption, continued, and extended use of radio frequency identification (RFID) technology: Australian Livestock Industry, Doctoral dissertation, Curtin University, Australia (2012)
20. Hsieh, J.J.P., Wang, W.: Explaining employee's extended use of complex information systems. European Journal of Information Systems 16(3), 216–227 (2007)
21. Lee, Y., Kozar, K.A., Larsen, K.: The technology acceptance model: Past, present, and future. Communications of the Association for Information Systems 12, 752–780 (2003)
22. Mathieson, K.: Predicting user intentions: comparing the technology acceptance model with the theory of planned behavior. Information Systems Research 2(3), 173–191 (1991)
23. Moon, J.W., Kim, Y.G.: Extending the TAM for a world wide web contex. Information & Management 38(4), 217–230 (2001)
24. Moore, G.C., Benbasat, I.: Development of an instrument to measure the perception of adopting an information technology innovation. Information Systems Research 2(3), 192–222 (1991)
25. Moberg, Å., Johansson, M., Finnveden, G., Jonsson, A.: Printed and tablet e-paper newspaper from an environmental perspective: A screening life cycle assessment. Environmental Impact Assessment Review 30, 177–191 (2010)
26. Mäntymäki, M.: Continuous use and purchasing behavior in social virtual world, Doctoral dissertation, Turku School of Economics, Turku, Finland (2011)
27. Nysveen, H., Pedersen, P.E., Thorbjørnsen, H.: Intentions to use mobile services: Antecedents and cross-service comparisons. Journal of the Academy of Marketing Science 33(3), 330–346 (2005a)
28. Nysveen, H., Pedersen, P.E., Thorbjornsen, H., Berthon, P.: Mobilizing the brand: the effects of mobile services on brand relationships and main channel use. Journal of Service Research 7(3), 257–276 (2005b)
29. Ottman, J., Stafford, E., Hartman, C.: Avoiding green marketing myopia: Ways to improve consumer appeal for environmentally preferable products. Environment: Science and Policy for Sustainable Development 48(5), 22–36 (2006)
30. Pelletier, L.G.: A motivational analysis of self-determination for pro-environmental behavior. In: Deci, E.L., Ryan, R.M. (eds.) Handbook of Self-Determination Research. The University of Rochester Press, NY (2002)
31. Podsakoff, P.M., MacKenzie, S.B., Lee, J., Podsakoff, N.P.: Common method biases in behavioral research: A critical review of the literature and recommended remedies. The Journal of Applied Psychology 88(5), 879–903 (2003)

32. Podsakoff, P.M., Organ, D.W.: Self-reports in organizational research: Problems and prospects. Journal of Management 12(2), 531–544 (1986)
33. Rodgers, S., Chen, Q.M.: Post-adoption attitudes to advertising on the internet. Journal of Advertising Research 42(5), 95–104 (2002)
34. Rogers, E.: Diffusion of Innovations. Free Press, New York (2005)
35. Scharl, A., Dickinger, A., Murphy, J.: Diffusion and success factors of mobile marketing. Electronic Commerce Research and Applications 4(2), 159–173 (2005)
36. Sharma, S., Mukerjee, S., Kumar, A., Dillon, W.R.: A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. Journal of Business Research 58(7), 935–943 (2005)
37. Spyridou, P., Veglis, A.: Exploring structural interactivity in online newspapers: A look at the Greek web landscape. First Monday 13(5), 1–10 (2008)
38. Steenkamp, J.-B.E.M., van Trijp, H.C.M.: The use of LISREL in validating marketing constructs. International Journal of Research in Marketing 8(4), 283–299 (1991)
39. Sussman, S.W., Siegal, W.S.: Information influence in organizations: An integrated approach to knowledge adoption. Information Systems Research 14(1), 47–65 (2003)
40. Tian, X., Martin, B., Deng, H.: The impact of digitization on business models for publishing: Some indicators from a research project. Journal of Systems and Information Technology 10(3), 232–250 (2008)
41. Trappey III, R.J., Woodside, A.G.: Consumer Responses to Interactive Advertising Campaigns - Coupling Short-Message-Service Direct Marketing and TV Commercials. Journal of Advertising Research 4(3), 382–401 (2005)
42. Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science 46(2), 186–204 (2000)
43. Venkatesh, V., Davis, F.D., Morris, M.G.: Dead or alive? The development, trajectory and future of technology adoption research. Journal of Association for Information Systems 8(4), 267–286 (2007)
44. Venkatesh, V., Thong, J.Y.L., Xu, X.: Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly 36(1), 157–178 (2012)
45. Wati, Y., Koo, C.: Toward green IS adoption behaviors: A self-determination perspective. In: Proceedings of 45th Hawaii International Conference on System Sciences (HICSS), Hawaii, USA (2012)

Appendix 1. Survey Items Used to Measure Constructs and Scaling

Eco-consciousness	I believe that using a tablet device is a good thing for the environment (X_1) I believe that using a tablet device would be an eco-conscious choice (X_2) I believe that using a tablet device would help to preserve natural resources (X_3) By using a tablet device, I believe I could help the environment (X_4)
Ease of use	I believe a tablet device is easy to use (X_5) I believe it is easy to get a tablet device to do what I want it to do (X_6) I believe my interaction with a tablet device would be clear and understandable (X_7) I believe that using a tablet device does not require a lot of my mental effort (X_8)
Extra services	Possibility to watch news videos (X_9) Get more information on certain factors in a news article by clicking the words (X_{10}) Possibility to see more picture material of certain news (X_{11}) Possibility to save interesting articles into an article bank on a tablet device (X_{12})
Social prestige	Using a tablet device would help me feel accepted (X_{13}) Using a tablet device would improve the way I am perceived (X_{14}) The use of a tablet device would give its owner social approval (X_{15})
Attitude	Using a tablet device would be interesting (Y_1) I would have fun using a tablet device (Y_2) The use of a tablet device would give me pleasure (Y_3)
Intention to use	Given that I had a tablet device, I predict I would use it (Y_4) Assuming a tablet device would be available for me I intend to use it (Y_5)
Intention to buy	I intend to purchase a tablet device in the short term (Y_6) I intend to purchase a tablet device in 5 years (Y_7)

The response options ranged from 1, “totally disagree” to 7, “totally agree”

Can Gamification Increase Consumer Engagement? A Qualitative Approach on a Green Case

Stavros Lounis, Xanthippi Neratzouli, and Katerina Pramatari

ELTRUN E-Business Research Center
Department of Management Science & Technology
Athens University of Economics and Business
{slounis,k.pramatari}@aueb.gr, xnera@dmst.aueb.gr

Abstract. The present study aims to identify the potential benefits deriving from the introduction of gamification elements in the consumer shopping process, in order to engage consumers in a more ecologically conscious behavior. Interviews with lead consumers show that the gamification of the shopping process results in increment of the stated intention to participate in the shopping process as well as an increment in the stated intention to purchase and an increase in the price premium consumers are willing to pay for environmentally friendly products.

Keywords: Gamification, Consumer Engagement, Price Premium, In-Depth Interviews.

1 Introduction

During the past few years Gamification has received the attention of the industry (Gartner Inc., 2011) as well as the academia (Deterding et al., 2011; Lee & Hammer, 2011; Simõesa et al., 2012). The phenomenon of its penetration in the corporate marketing strategy and the way the practitioners have found it to elicit and enhance user engagement though, has not been adequately studied in the academic literature. In the case that we examine consumers in particular, gamification literature is virtually inexistent. In an ever-changing scenery of consumer consumption behaviors, the effect of sustainability has risen in various industry sectors. Suppliers and retailers of fast moving consumer goods focus their efforts on producing environmentally friendly products and in the latter years efforts are turning to collaboration on achieving the goal. Although the development of sustainable consumption patterns and practices has been discussed in the academic literature since the 60's (Packard, 1960), until the present time the selection of products is not heavily influenced by the environmental impact the products have, rather than by other factors such as price, convenience etc. (Gaspar & Antunes, 2011; Gadenne et al., 2011; Faiers, Cook, & Neame, 2007). That being the case, the question remains: How will the collaborative ecological efforts of the industry to reduce carbon emissions throughout the supply chain be extended to the consumer? More importantly can consumers endorse sustainable patterns of consumption and become green consumers via the employment of gamification in the

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 200–212, 2013.

shopping process. The aforementioned question is the motivation behind the present research. In the next section we provide an overview of the existing literature on Gamification and current research findings on the green consumer. The third section presents the methodology of the study and the fourth and fifth section present the findings and discussion of the results as well as the academic and managerial implications. The final sections refers to the limitations and directions for future research of the present study.

2 Literature Review

2.1 Gamification

In the academic literature gamification is starting to gain momentum and different research streams employ different definitions in relation to the point of view they examine it and its effects. Emphasizing on the overall goal of gamification from the service marketing perspective, Huotari & Hamari (2012) define gamification as "a process of enhancing a service with affordances for gameful experiences in order to support a user's overall value creation". Another definition comes from Deterding et al. (2011) who identify gamification as "… the use of design elements characteristic for games in a non-game context". That approach presents a more systemic view of the term, pointing out the process of decomposing games into building blocks and introducing them into areas that can benefit and didn't previously employ such techniques. A different view of gamification presented by Zichermann and Cunningham (2011) endorses gamification as a "process of game-thinking and game mechanics to engage users and solve problems". All aforementioned definintions have merit in the way they address gamification. However, at present time, a universally applicable definition of gamification is not extant. For the purposes of this research we extend the definition of Gamification presented by Zichermann as follows:

Gamification is the Process of game-thinking and game-mechanics to engage the consumer in the non-gaming context of shopping in order to drive engagement and enhance the process of behavioral shift.

2.2 Green Consumer and Shopping Behavior

Having focused on the consumer and the potential shift towards sustainable consumption practices via the employment of gamification, we proceed to examine the attributes that formulate the profile and behavioral patterns of the Green consumer. These attributes have been studied extensively in the literature both from the individual consumer perspective (Straughan & Roberts, 1999; Roberts & Bacon, 1997) as well as at an aggregate national level (Dunlap, Mertig, & E., 2000). Pertaining to the individual consumer perspective research in different industry sectors reveals that although environmental factors are considered at the point of purchase, they are not the sole or most important factors affecting consumer choice (Gaspar & Antunes, 2011; Gadenne et al., 2011; Faiers, Cook, & Neame, 2007). However,

Straughan & Roberts (1999) identified that the green consumer behavior is driven by the person's "…belief that individuals can play an important role in combating environmental destruction…", that is consistent with the findings of the stream of research that explores the psychological benefits that arise from the person's contribution to the ecological common good (Wiser, 1998; R. & M., 2006; Hartmann & Apaolaza-Ibáñez, 2012). The aforementioned stream of researchers portrays a consumer driven by internal motives on their choices where as the current practice in the FMCG sector is focused on incentivising the consumer with external redemable rewards via loyalty programs (Meyer-Waarden, 2008; Smith & Sparks, 2009). That type of incentives is found to have results under the premise that the loyalty program is ongoing, yet problems arise in its termination (Yi & Jeon, 2003; Leenheer et al., 2007; Meyer-Waarden & Benavent, 2009). The missing link in that perticular problem, caused by the short term effect of loyalty programs, could be addressed by the introduction of a mechanism that enhances the process and the incentive mechanisms that support psychological benefits in parallel to monetary external rewards, namely gamification.

2.3 Gamification as Means to an End

Various industry sectors have identified the potential for increment in consumer interactions and jumped on the gamification wagon for the ride. Retailers in India have identified it as an upcoming trend of social media marketing and their customer centric initiatives include gamification in their core process in order to "drive engagement and participation" (Archana, 2012). In Education, gamification has been found to have great potential to motivate students (Lee & Hammer, 2011; Simõesa, Díaz Redondob, & Fernández Vilasb, 2012). In the sustainability sector in particular, research conducted by Kuntz et al. (2012) resulted in the introduction of gamification in the sustainability awareness and efforts of individuals had positive outcome in saving energy, water and reducing gasoline use. As with the afforementioned sectors that have benefited from the introduction of gamification, we aim to examine the potential for benefit in the context of FMCG and the green consumer. The Fast Moving Consumer Goods sector comprises of businesses that offer products produced for frequent consumption as: Food and beverages, houshold goods, sports goods, personal care and cosmetics etc.

3 Qualitative Research

In order to gain insights into consumer reactions to the introduction of gamification in the shopping process in the FMCG sector, in-depth interviews with fifteen consumers were conducted (47% male). The focus of the in-depth interviews was twofold: First, to garner consumer perceptions on green consumption in the FMCG sector and second to identify potential benefits from the introduction of gamification. Questions were also asked in regards to the general shopping behavior and patterns of the individuals.

The method of consumer selection was based on Matthing, Kristensson, Gustafsson, & Parasuraman's (2006) typology of consumers. In this typology, authors divide consumers in 6 categories: leaders-explorers, pioneers, skeptics, paranoids and laggards, based on the characteristic of how innovative a person is; part of the characteristic of innovation is "..the tendency of people to look for different aspects of the reality" (Matthing et al, 2006). For the present research purposes only people who belong in the categories of leaders-explorers, pioneers and skeptics were selected for participation. The identification of lead consumers was performed via an offline questionnaire prior to the in-depth interviews.

Additionally, in order to determine the degree of ecological worldview and ecological conscious consumer behavior of the participants and further segment them, two additional offline questionnaires were administered. The first was the New Environmental Paradigm – NEP (Dunlap, Van Liere, Mertig, & Jones, 2000) and the second was the Ecological Conscious Consumer Behavior (Roberts, 1996). The latter in particular was utilized to measure the extent to which the respondents purchase goods / services that they believe have a "more positive" impact on the environment in relation to their counterpart alternatives. The consumers selected for the in-depth interviews were screened and assessed to have the lead consumer profile as well as various degrees of ecologically conscious consumer behavior (ECCB) and ecological worldview (NEP). The resulting sample (Figure 1) of lead consumers consisted of both male (47 %) and female (53 %) consumers. The age range of the sample was segmented into [18-24 at 33 %, 25-29 at 40 % and 30-45 at 27 %] and the group was further segmented based on their environment-conscious behavior into two segments (high ECCB 55.5 % and medium/low ECCB 44.5 %) in order to identify correlations of gamification application on various levels of the green consumer.

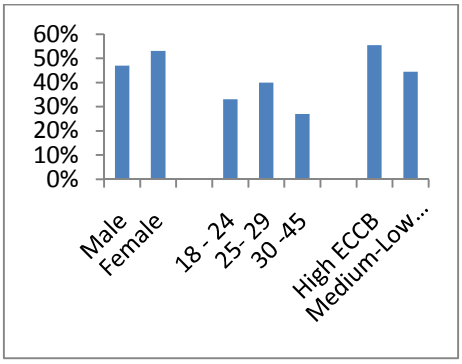


Fig. 1. Demographics of lead consumer sample

Although the number of the consumer sample consisted of fifteen consumers, with careful sampling and thorough collection technique, a small amount of in-depth interviews can result with data capable of addressing the research question (Holloway, 1997). The researchers continued to sample until there was no production of any new information or insights following the theoretical saturation general rule of qualitative research. In the present study theoretical saturation was deemed to be achieved and satisfactory for valid sampling after fifteen consumers. The present study is consistent with the suggested valid range of case sampling of more than ten cases (Eisenhardt, 1989). The information gathered during the interviews was content analyzed and coded by two independent coders into categories pertaining to the present research. A number of categories were decided upon ahead of time (e.g. Game Mechanics for incorporation, Shopping process) while other

categories were identified based on the lead consumer responses (e.g. Flow of gamified service). In the case of disagreement on the classification of any particular statement, the disagreements were resolved upon joint discussion.

4 In-Depth Interviews Results

The in-depth interviews results address three major categories: (a) the current and ecological shopping behavior (b) services to support the shopping process / ecologically conscious consumer behavior and (c) gamification insights on driving consumers in the common effort to endorse sustainable consumption. The in-depth interviews results follow:

4.1 Current and Ecological Shopping Behavior

The lead consumers selected for the in-depth interviews presented a near and above average environmental worldview as identified by the initial screening process and were further on divided into the consumers that presented high ecologically conscious consumer behavior and medium to low ecologically conscious consumer behavior (ECCB) in order to classify shopping behavior and insights respectively.

The lead consumers classified as high ECCB display higher tendencies to request and even search for information pertaining to the environment-friendly products / enterprises and factor the results to their product selection criteria. As a valuable insight, the majority of the consumers stated that when they are presented with eco-relevant product information they hold accountable the product on its claim and if found that it is valid (as perceived by their definition of validity) are willing to pay a 20 % price premium to purchase the product on average (reported up to 30 %). However, if they find the claims to be invalid and not substantial, this results in considering it as "green-washing" and the product is not eligible for purchase.

The lead consumers classified as medium-low ECCB display lower tendencies to search for eco-friendly products and factor that in their purchasing decision at a lower factor. However, they still consider it a factor. A valuable insight on the specific type of consumers is that the majority stated that although in the past they were not shopping ecologically, they start (at varying rate) to turn towards that as well. In this category, the acceptable price premium for environment-friendly products is 10 % on average (reported up to 25 %).

Product Ecological Friendliness Characteristics and Perceived Quality

The majority of consumers, when asked if they perceive an ecologically-friendly product to have inferior characteristics (such as taste for food, cleaning efficiency for detergents etc.) when compared and contrasted to another product (not claiming eco-friendliness) responded that they strongly believed it to be of equal if not higher quality. Respondents even claimed that "nowadays technology has evolved to such degree that products can be of quality and be eco-friendly at the same time". Additionally,

when presented with price differences amongst eco-friendly and regular products the respondents stated that they would have no problem paying a premium price to buy products that are eco-friendly. Specifically, 61 % of the total respondents are willing to pay up to 10 % premium price for the eco-friendly product, 26 % are willing to pay up to 25 % and finally 13 % are willing to pay up to 30 % (Figure 2), indicating that they value the ecological perspective of the products and understand that the investments required for a product to become eco-friendly are of a certain magnitude.

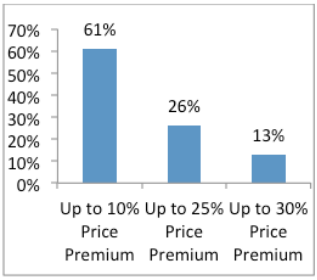


Fig. 2. Willingness to pay price premium for eco-friendly product

Ecological Friendliness Product Information Placement

The majority of the respondents could identify a number of product categories where they are loyal to specific brands. Lead consumers, identified as "able to select a product taking into consideration factors other than brand loyalty", were then presented with a three-step scenario relating to the product selection process in the respective categories. In that scenario they stated their process of thinking. The first step is that they observe a product having a label with a message claiming eco-friendliness and sustainability. Their response was that they are not convinced and they request additional information substantiating the claim at a 74% and "probably is environmental" at a 26 %. The second step requested input from the lead consumers on the information they would like to be presented with. From the total of the respondents 87% indicated that they would like to see some form of substantiation of the eco-friendly claim of the label and of those respondents 33% preferred data as facts where 53% preferred a visualization of the data / facts (the remaining did not state a clear selection). The final step was the option to have the aforementioned data on the shelf by the product price tag where 53% consider it positive and stated that they would read it and consider the environmental factor higher and 20% consider it negative (described as green-washing). The remaining did not state a clear selection. From the second step on, the majority (74%) of the respondents described the eco-friendly product claim to be of substance. A valuable insight from the scenario was that lead consumers identified and pointed out the necessity for standardization of ecological measures and claims. In the process of the interviews, all respondents stated that they lack the information pertaining to the ecological footprint each and all products have. Additionally, as a measure to prevent themselves from mal-practices the majority pointed out the need for some form of standardization of the measuring and reporting of the ecological footprint of the products within a retail store. This unified measure should be obligatory, span across all product categories and address the entirety of the product lifecycle from creation to disposal. "If it's (the measure) not the same how can I compare them? And If I can't compare them how do I know which is better?" and "Products have nutritional information on them and I see which is closer to my needs. That is something I want for the environment as well" are responses that illustrate the perception. Furthermore, a number of consumers presented the energy class of the

electrical appliances as a form of elaboration on the necessity of standardization in all sectors.

4.2 Services to Support the Shopping Process / Ecologically Conscious Consumer Behaviour

The majority of the lead consumers do not regularly use online technologies to support their shopping needs with exceptions being the occasional use of Internet for product-related information search and notes on a smartphone about the products for purchasing. The services identified by consumers related to the environmental profile of products and to the general shopping mission are presented in the Table 1:

Table 1. Services related to the ecological profile of products and shopping process

Service Supporting…	Service features and consumer endorsement
1. Identification of shopping needs	Monitor stock levels of household products, maintain historic of product needs, calculate consumption volumes / product (Stated by 46% of the consumer sample)
2. Shopping needs recommendation	Optimization of the product purchase selection based on budget and cost scenarios (Stated by 60% of the consumer sample)
3. Product shopping planning	Plan upcoming product purchases, retrieval of current product prices (Stated by 67% of the consumer sample)
4. Shopping list	Online / shared by household members. Available in both web and mobile versions. Consumer "Assignment to shop specific products" capabilities (Stated by 87% of the consumer sample)
5. Retail store selection	Retrieval and Comparison of product price / availability between proposed retail stores (Stated by 46% of the consumer sample
6. Product reviews	Product information and reviews from other consumers (Stated by 46% of the consumer sample)
7. Product Information (excluding ecological related product info)	Product information, personalized or aggregate product sales promotion, complementary information (cooking time etc.) (Stated by 66% of the consumer sample)
8. Product purchase selection	Barcode / QRCode / NFC identification of consumer selected products, Shopping basket cost (Stated by 73% of the consumer sample)
9. Self checkout / Electronic payment	Ability to self checkout and electronic payment (Stated by 60% and 33% of the consumer sample respectively)
10. Online loyalty points	Online assignment, presentation and exploitation of the loyalty programs the consumer participates in (Stated by 40% of the consumer sample)
11. Online shopping/delivery	Ability to shop online / Delivery (Stated by 40% of the consumer sample)

Extending the Services to Support the Green Consumer

As stated by the majority of the lead consumers in the course of the in-depth interviews, the aforementioned services pertaining to the need for identification and selection of product for purchasing [Services: 1-4, 6,7] should "include information on ecologically related information" such as CO_2, energy efficiency, recycling

information etc. Further on, the majority of the lead consumers proposed an eco-awarding program as an extension of the online loyalty points program [Service 10] in order to give them incentives to select ecologically friendly products prior and during the moment of purchase. This includes:

12. Service of product "eco-friendliness": The consumers requested a service that would inform them on the sustainability and eco-friendliness of the product. The service should provide information related to the production and distribution process followed for the product, its certifications, eco-comparative product information etc. (Stated by 73% of the consumer sample for extension of services 1-4, 6,7).

13. Awarding program service: The proposed service would include a "loyalty program" type of service where ecologically conscious consumer behavior would be awarded. The consumers described a mechanism where the consumer collects points (similar to the loyalty points, e.g. ECOPoints, ecology points, green points, green bonus etc.) when one purchases eco-friendly products and later on he/she can exchange the received points for price reduction or free products. The presentation of the proposed service was described being available both online and offline (Stated by 87% of the consumers consumer sample as an extension of service 10). The following table presents the services that have been identified by lead consumers during the in-depth interviews. The percentage indicates the consumers that identified the need for each respective service during the interview.

4.3 Gamification Insights on Driving Consumers to Endorse Sustainable Consumption

In the process of the in-depth interviews, consumers expressed their opinions on the various processes and stated their terms and subsequent degrees of engagement and expected benefits from participation in such an effort. The lead consumers identified the following in terms of gamification elements eligible for introduction in the ecological FMCG shopping context. Consumers indicated that in the process of creating an experience that would drive themselves as well as other consumers to behave in an ecologically conscious manner in the shopping context, a connection to the actual shopping process is important. A direct link between the shopping choices and the game mechanics should exist. The following statement illustrates the proposal: "When you shop in an ecological (or not) way something is triggered and something happens". This direct link from shopping behavior to a game-like experience (or actual game) could be established (as described by the majority of the consumers) by the introduction of a mechanism "loosely related to the loyalty scheme that already exists". At present, when consumers shop in a specific manner, they receive loyalty points that they later on redeem for various commodities or price reduction. In extension to the previous concept, the consumers described a service where you are presented with the option to collect points (termed ECO Points, Environment Points Green Points, Nature Points etc.) when you opt to purchase ecologically friendly products. In regards to the point awarding system, 93 % of the consumer sample stated the point-awarding mechanism should be positive. Elaborating on the previous, when a consumer selects to purchase a product that is "eco-friendly" he / she should

receive points that add up to a (or a set of) accumulated total(s) towards a certain and clear goal.

Extending the point-awarding system and transferring it into a possible game (or game-like schema), consumers identified the key characteristic of comparison. Operating under the assumption that shopping in an ecologically conscious way leads to the accumulation of points, consumers stated interest in the ability to see the impact of their choices compared to the impact of other consumers' choices.

Levels of comparison: The lead consumers indicated different setups of comparison as optimal for engagement efficiency. In the context of comparison with their close friends in the form of ranking, 67% of the sample stated that they would like to know their own environmental consumption (past, current and evolution of) as individuals as well as where they rank amongst their friends. As the comparison expanded to include acquaintances, the percentage of consumer's interest to receive comparative results dropped to 33 %. A further extension to the degree of city improved the interest to receive comparative results to 40 %, and the extension within the country presented a 26% of interest to receive comparative results. From then on, comparison on the level of continent as individuals was described as out of context and the participation intent was reduced. Upon concluding the comparison levels, the majority of the participants stated that besides individual comparison they would be interested in a community approach of the same design.

Community formation and comparison: Pertaining to stated intent of participation, the prevailing community formation design for comparison and ranking (based on ecological consumption and behavior) were identified as the City Community, Country Community and Virtual Community. (a) City community: The respondents expressed increased interest in formation of the City Community for degree of community comparison especially when the incentives for participation spanned from the virtual world to the real world and were of environmental nature. The participation to the game and engagement in environmentally-conscious shopping behavior presented a 15% average of acceptable raise to the premium price for eco-products based on incentives and during a contest. (b) Country community: The country community, although presented as an important driver towards the intent to participate, did not display as high engagement in ecologically-conscious shopping behavior as the participants felt that the incentives to participate in a contest were out of reach. (c) Virtual Community: Besides physical communities, consumers displayed interest in the possibility to create custom and virtual communities and participate in contests mostly bearing virtual prizes.

Incentives for Participation

Another important gamification element is the incentive mechanism(s) employed. As the point-awarding system presents the common comparison denominator the incentive mechanisms are to employ it. During the in-depth interviews, various incentive mechanisms were proposed with the most prominent being: (a) Monetary incentives (consumer / environment): After the accumulation of a defined amount of points (absolute or relative to time / effort) the consumer would be awarded with price reduction on ecological friendly (or not) products of their choice. The incentive was found

applicable to individual rewarding (contests or continued) and is supported by 67% of consumers. Consumers additionally stated a different type of monetary incentive mechanism where the reward would be in the form of "Doing something for the environment". In this case, consumers stated that they would participate in the context of a contest if the defined prize addressed an environmental issue / supported a cause. The aforementioned type of rewarding systems was found applicable to community rewarding (mostly contests) supported by 67% of consumers. (b) Virtual incentives: The incentive mechanism was described as extendable / transferable to the virtual / online / game world. This type of incentives would include (but not limited to) the awarding of personal and community badges (eco stars, trees, virtual forests etc.) as defined by the respective rules of the game(s) or game-like application, the customization of digital consumer representations (avatars, profiles) and virtual environmental quests (both educational and entertainment nature). The respondents indicated that although monetary incentives are important, virtual incentives are sought out particularly for continuous engagement with the process in its entirety (in-between the "real world" incentives) as supported by 73 % of consumers. (c) Social media incentives: Another form of incentive mechanism portrayed by the majority of the lead consumers was the incorporation of social media in the process. This particular mechanism, although adopted by the majority of the sample, was adopted at various steps of the gamified experience and in varying degrees. Elaborating on the previous, the consumers stated that their social media presence and exposure during the gamified experience should be under their control.

Off-line approach: The aforementioned gamified experience was found to be transferable to an offline context as well by the consumers, assuming that the retailer/supplier would provide them with the appropriate infrastructure at the retail store. Although the point of contact to the service would be limited to the time the consumer was within the store, the experience was equally valued by the majority of consumers and 53 % stated that they felt that they would still participated.

5 Discussion - Implications

As the focus of the in-depth interviews was to garner consumer perceptions on gamification and green consumption in the FMCG sector and to identify potential benefits from the introduction of gamification, results indicate a potential benefit from the introduction of gamification in the FMCG process. The percentage of consumers that stated their intent to participate in the gamified process presents an opportunity for the sector to exploit a new marketing medium in their strategies. The introduction of gamification though is not panacea in consumer engagement. The in-depth interview results indicate that not all consumers respond / endorse all aspects of gamification in the same degree and different gamification dimensions, result in different outcomes in terms of engagement / stated intention to participate etc. Thus the need for personalization. The ability of the proposed gamification scheme to be customizable and personalized should span throughout the entirety of the selected and implemented processes to become efficient. Although the previous is a generally applicable

condition, the interviews showed consumer clusters of common beliefs and behavioral patterns creating thus an initial dataset of common attributes. An important insight pertains to the issues that arise from the consumer privacy concerns in relation to the need for personalization. In the process of participation in the aforementioned gamified scenarios, the consumers stated that the quality / amount of information available for processing and feedback should be subjected to regulations securing them from misuse or alternative use (other than the gamified service itself). The previous insights comprise the academic implications of the present research.

An additional important insight pertaining to the efficiency of the engagement provided by the utilization of gamification to the ecologically conscious consumer behavior and its extended managerial implications is presented by the stated behavioral conversion whilst participating in an environmental gamified contest. In a scenario of their choice, the lead consumers stated their intent to (a) pay premium price for eco-friendly products at a higher margin and (b) select to purchase eco-friendly products on more occasions and product categories. The stated behavioral conversion was higher (and maintained) throughout the existence of the game: 80% stated that the acceptable premium price could rise from 5% to 25% relative to their initially stated maximum acceptable premium (15% on average). In the case of the removal of the gamified service(s), 33% of the consumers stated that they would be more (in comparison to their previous state) ecologically conscious on their shopping behavior. The latter illustrates the potential for shift in shopping practices towards more sustainable practices. In conclusion the results deriving from the analysis of the in-depth interviews, support the potential benefits that derive from the introduction of gamification in the FCMG sector in regards to the ecologically conscious consumer behavior.

6 Limitations – Research Agenda

This research is not without limitations. A potential limitation is the use of student respondents that although are consumers and the FMCG shopping context was deemed relevant, the homogeneous nature of the sample (education) likely resulted in a more restricted variation of personal characteristics than if a more heterogeneous group was used. Another possible limitation arises from the inherent limitations of in-depth interviews as the probability that the interviewee may distort information through recall error, selective perceptions and desire to please interviewer, yet all appropriate measures was set in order to negate the limitations of the in-depth interviews (Lofland & Lofland, 1995; Patton, 1990).

The present article presents the qualitative analysis and the consumer insights that will be used as input on the research agenda. Future research includes a lab experiment where the applicability of gamification will be examined in a simulated shopping environment that employs gamification processes and the analysis of results in correlation to extant consumer behaviour models, a European survey to test the appropriate game mechanics in consumers from different countries and a filed study to measure the impact of the introduction of gamification to the consumer's endorsement

of sustainable patterns of consumption and the transition to a green consumer. Although gamification is in its infancy, the future seems prominent.

Acknowledgments. This work has been supported be the e-SAVE research project, co-funded by the European Commission (FP7-ICT-288585).

References

1. Archana, B.: A Case Study on Retailing in India. EXCEL International Journal of Multidisciplinary Management Studies 2(7), 178–188 (2012)
2. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From Games Design Elements to Gamefulness: Defining "Gamification". In: Proceedings of The 15th International Academic Mindtrek Conference, Tampere, Finland (2011)
3. Dunlap, R.E., Mertig, A.G., Jones, R.E.: Measuring Endorsement of the New Ecological Paradigm: A Revised Nep Scale. Journal of Social Issues 56(3), 425–442 (2000)
4. Dunlap, R., Van Liere, K., Mertig, A., Jones, R.: Measuring Endorsement of the New Ecological Paradigm: A Revised NEP Scale. Journal of Social Issues 56, 425–442 (2000)
5. Eisenhardt, K.: Building Theory from case study research. Academy of Management Review 14(4), 592–611 (1989)
6. Faiers, A., Cook, M., Neame, C.: Towards a contemporary approach for understanding consumer behaviour in the context of domestic energy use. Energy Policy 35, 4381–43390 (2007)
7. Gadenne, D., Sharma, B., Kerr, D., Smith, T.: The influence of consumers' environmental beliefs and attitudes on energy saving behavious. Energy Policy 39, 7684–7694 (2011)
8. Gartner Inc. (May 9-10, 2011) from Gartner Inc., http://www.gartner.com/it/page.jsp?id=1629214 (retrieved January 12, 2012)
9. Gaspar, R., Antunes, D.: Energy efficiency and appliance purchases in Europe: Consumer profiles and choise determinants. Energy Policy 39, 7335–7346 (2011)
10. Hartmann, P., Apaolaza-Ibáñez, V.: Consumer attitude and purchase intention toward green energy brands: The roles of psychological benefits and environmental concern. Journal of Business Research 65, 1254–1263 (2012)
11. Holloway, I.: Basic concepts for qualitative research. Wiley-Blackwell (1997)
12. Huotari, K., Hamari, J.: Defining Gamification - A Service Marketing Perspective. In: Proceedings of the 16th International Academic MindTrek Conference. Tampere, Finland (2012)
13. Kuntz, K., Shukla, R., Bensch, I.: How Many Points for That? A Game-Based Approach to Environmental Sustainability. In: ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA, pp. 126–137 (2012)
14. Lee, J., Hammer, J.: Gamification in education: What, how, why bother? Academic Exchange Quarterly 15(2) (2011)
15. Lofland, J., Lofland, L.: Analyzing Social Settings: A Guide to Qualitative Observation and Analysis, 3rd edn. Wadsworth, Belmont (1995)
16. Matthing, J., Kristensson, P., Gustafsson, A., Parasuraman, A.: Developing successful technology-based services: the issue of identifying and involving innovative users. Journal of Services Marketing 20(5), 288–297 (2006)

17. Packard, V.: The Waste Makers, pp. 184–202, pp. 255–263. Penguin Harmondsworth (1960)
18. Patton, M.: Qualitative Evaluation and Research Methods, 2nd edn. Sage, Newbury Park (1990)
19. Wüstenhagen, R., Bilharz, M.: Green energy market development in Germany: effective public policy and emerging customer demand. Energy Policy 34(13), 1681–1696 (2006)
20. Roberts, J.: Green consumers in the 1990s: profile and implications for advertising. Journal of Business Research 36(3), 217–231 (1996)
21. Simõesa, J., Díaz Redondob, R., Fernández Vilasb, A.: A social gamification framework for a K-6 learning platform. Computers in Human Behavior (2012)
22. Straughan, R.D., Roberts, J.A.: Environmental segmentation alternatives: a look at green consumer behavior in the new millennium. Journal of Consumer Marketing 16(6), 558–575 (1999)
23. Wiser, R.: Green power marketing: increasing customer demand for renewable energy. Utilities Policy 7(2), 107–119 (1998)
24. Zichermann, G., Cunningham, C.: Gamification by Design Implementing Game Mechanics in Web and Mobile Apps. O'Reilly Media, Inc., Canada (2011)

Designing to Promote a New Social Affordance for Energy Consumption

Lara S.G. Piccolo, Heiko Hornung, Cecília Baranauskas, and Roberto Pereira

Institute of Computing, The State University of Campinas - UNICAMP
larapicc@gmail.com, heix@gmx.com,
cecilia@ic.unicamp.br, roberto.ihc@gmail.com

Abstract. The relation between energy consumption and its environmental impact is weak or even almost inexistent in the Brazilian scenario due to sociocultural reasons. This work takes advantage of an experimental smart grid deployment scenario to propose the concept of Social Eco-feedback Technology. It aims at establishing this relation by promoting a new social affordance (behavior pattern), favoring a new way of perceiving energy. The concept of affordance and artifacts of the Organizational Semiotics are extended to encompass both the analysis of actual and intended behavior through a new technology design.

Keywords: Eco-feedback technology, Energy consumption, Affordances, Motivation, Human-Computer Interaction, Organizational Semiotics.

1 Introduction

The *smart grid* wave [36] is a manifestation of the need to rethink the way energy has been generated, distributed, and used around the world, especially considering that the natural resources of the planet have already been over-consumed, although not equally distributed among the whole population [37]. Energy companies throughout the world have specific interests in deploying smart grid technologies [4] capable of lowering costs of energy distribution, reducing demand during peak time, etc. Mainly in developed countries, providing more detailed information about how energy has been consumed, a possible feature that might be related to the smart grid technology, has been considered a crucial step for consumption reduction as a response to the pressure for restraining the climate change [4].

In the Human-Computer Interaction (HCI) domain, energy consumption is emerging as an important topic of interest [23] and many different approaches have been adopted to design consumption displays intending to promote a more sustainable individual behavior. According to the review of energy related-work by Pierce and Paulos [24], other studies that focus on the design and evaluation of energy consumption feedback are situated in domestic upper-middle class population of "developed" contexts.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 213–225, 2013.

Some particularities of Brazil have framed energy-related research in the country in the last years, in such way that the importance of the connection between consumption and the natural environment by individuals has been underestimated.

Comparing it to other countries, especially to the United States and Europe, the energy is mostly generated by hydro plants, which are less polluting than thermoelectric power, commonly used worldwide [7]; the energy consumption per capita in Brazil is below the world average – in 2009 the Brazilian average reached 2.206 kWh, while the world average was 2.806 kWh; in the United States it was 12.900 kWh and in Europe 6.063 kWh [2].

Nevertheless, Brazil was set in the last position in a 17 countries ranking of a survey that inquired people about the correlation between environmental impact and energy consumption. Asked about factors with a negative impact on the environment, only 27% of Brazilians respondents mentioned individual energy consumption [1]. Results from a survey situated in this research scenario reinforce this panorama. Participants were asked "how do you believe that saving energy might contribute to the environment?" In low-income areas, 85% of respondents said to not know the answer, while only 7% correlated energy usage with new dams or power plant building. These results are related to socioeconomic level of the population, but the "I do not know" answer was in general predominant [22]. The survey, part of two smart grid programs also pointed out the consensual interest (95%) in receiving more information about conscious consumption, as well as the lack of awareness of people's own consumption: in one of the areas, 90% declared not to know the average of their consumption in kWh.

While these numbers provide an overview about how people perceive individual energy consumption and connect it to environmental impacts, 71 new hydro plants have been planned and built from 2008 to 2017 according to the official national plan of energy generation [3], in addition to the 140 already in operation, in order to supply the increasing demand resulting from the economic development and the forecasted population growth.

Coping with this scenario, this study aims at promoting a new social affordance [30][17] regarding energy consumption, which means favoring a new way of perceiving and relating to energy and the natural environment, leading to different patterns of behavior. In line with [4] and [24], this study intends to take advantage of an experimental deployment of smart grid in Brazil to establish this new relation by means of a technology introduction that has been designed with this purpose. To situate this investigation in the Brazilian context considering socio and cultural aspects in the design, the Organizational Semiotics (OS) [30][17] approach has been the theoretical basis.

In the next session, an overview eco-feedback technology is presented highlighting the need to consider the social context of its usage. Then, we propose the Social Energy Eco-Feedback Technology concept, by mapping the reality into an ontology chart. We also propose a new approach to use this artifact for representing intended behavior with the new technology design. Further session describes norms that guide the technical device development, followed by a discussion regarding OS methods and results. The last session concludes and points out future works.

2 The Role of Eco-feedback Technology

In the energy consumption domain, *feedback* has been defined by Spagnolli [29] as "information about the consequences of household actions that involve electricity consumption", and *awareness* as a condition for conservation, since it represents the "knowledge users acquire about how and why to reduce waste by operating devices more efficiently, including hints that show a larger environmental impact".

Ambient displays, mobile devices, and online visualizations are possible types of Eco-feedback Technology (EFT) designed to incite the transformation of individual's behavior regarding the environment [9]. The presence of an EFT was found to lead to a consumption reduction between 5 and 20% [4][23].

Froehlich et al. [9][8] evaluated design aspects and motivational techniques for EFT such as goal-setting, comparison, incentive, etc., pointing out some open issues in the HCI area, especially when considering the social domain of the EFT usage, such as the effectiveness of competition as a motivational strategy. In [24], the authors identified works that have raised some criticism regarding the approach that deals only with individual's behaviors, disregarding the social dynamics. Dourish [6] argues that this dominant approach into environmental topics in HCI is self-limiting. Besides addressing cultural and political aspects in the design, he suggests to "connect people *through* their actions and their consequences", persuading people by the empowerment of collective actions, instead of individually connecting actions to their consequences.

Other studies are also based on social strategies: [15] evaluated the effectiveness of competition, which yielded controversial results regarding this strategy; [5] proposed motivating social environments as one of the mechanisms to involve consumers with the feedback technology. This strategy relies on Social Norms and *social proofs*, which consider that people act in a certain way to be in line with action of others in similar context [12].

Hall [13] argues that a technical device is the best way to trigger changes in peoples' behavior. Rokeach [26], and Sanders and Atwood [28] present different experiments where computer and mass media are used to let people become aware of their values. They assert that when individuals become aware of the contradictions between their conceptions of self and their values, attitudes, or behavior, they will reorganize their values and attitudes, and thus, their behavior, in order to make them more consistent with their conceptions of self.

In the next session, a social eco-feedback technology is presented, starting from the concept of affordance, which supports the proposal of a technical solution.

2.1 Social Affordances and Eco-feedback Technology

Affordance is a concept associated to how people relate to the environment and to the things that are part of it. Gibson [11] defined affordance in the ecological context as a result of the relationship between physical properties of the environment and the experience of the perceiving actor, which is subjective. Norman applied this concept to design [20] proposing later [18] a distinction between *real affordances* — related to

the physical properties of the world — and *perceived affordances*, which instead of that are subjective representations in the mind. Those properties determine how the thing could possibly be used. Taking advantage of it, no labels or instructions would be needed. More recently [19], Norman reviewed the concept centering it in the social domain. *Signifier* is the term Norman proposes to replace affordance, meaning what people need, and what design must provide. *Social signifier* is the most important class of signifier because most actions people do are social. For him, social signifiers are broader and richer than affordaces. They include culture and experiences, similarly to Stamper's social affordance idea [17].

Stamper [10] extends Gibson's concept of affordances of the ecological environment to the social environment, arguing that people as agents acting in the social environment are dependent on the knowledge that has been built up and handed down from generation to generation in a society, subsidizing the idea of *Social affordances* as repertories of behavior tuned to the social environment.

Darby [4] applied Gibson's concept of affordance to understand how householders have perceived and used different types of consumption feedback considering the energy bills and meters, the in-home displays, and on-line visualizations. According to Darby, introducing affordances in the study of the new technology centered the smart metering technology development on the users and in their relation to energy and the new devices.

The concept of social affordance and its nuances in terms of how people relate to the environment and socially interact with it influences the proposal of a technical artifact that aims at changing the way people perceive and relate energy and the natural environment to their actions.

3 A Social Energy Eco-feedback Technology

Differently from most ETF proposals, which are contextualized in the domestic domain [24], this study proposes an EFT that aims at tackling energy consumption collectively, within a social group. The Social Energy Eco-feedback Technology (SEET), or SEETree, due to the metaphor of a tree applied as a signal to connect energy and the natural environment, consists on a public display and a physical installation located in gathering areas to configure and represent results of energy-saving collective actions. The way the tree is lightened represents achievements of a collective saving action. The SEETree design concept is grounded in the Organizational Semiotics approach as follows.

Organizational Semiotics (OS) is a discipline that studies information and its functions in organized domains, such as a company, a digital system, or the introduction of a new technology in society, as is the focus of this study. SO provides methods and techniques for understanding and modeling information systems, considering social and human activities as part of this system [17][33].

A set of OS methods named MEASUR – Methods for Eliciting, Analyzing and Specifying User Requirements [32][34] – supports requirements analysis by considering that an organization is a system of norms that the members share through

signs (information). A norm then is a field of force that influences the members of the community with respect to how they behave or think [35] and can be understood as 'pattern' or 'standard', governing meanings, intentions, knowledge, responsibilities and influence exerted.

In order to propose an information system that intends to promote individual and collective awareness of electricity usage, this study relies on two methods of MEASUR: (i) Semantic Analysis Method – SAM [16]: proposes semantic models of patterns of behaviors (shared meanings) based on agents and affordances, to map ontological dependencies among them; (ii) Norm Analysis Methodology – NAM [34][33][17]: models the social rules identified and specified as the conditions and constraints for realizations of the affordances, considering typical behavior of research scenario gathered by surveys [22].

3.1 SAM and Social Affordances

We subscribe to the notion that the world is socially and subjectively constructed [30]. Considering this, Stamper proposed that Gibson's theory of affordance can be naturally extended to the social world for studying social behavior, and introduced the concept of social affordance, which has been applied in the OS domain as "a pattern of behavior that is shared in a social community", has a start and finish time, and a starting and finishing authority [10][17].

In this sense, before promoting a change in the way people relate to energy, it is important to understand current patterns of behavior, which are, according to literature, mostly unconscious and guided by the environment [14][25]; and according to data gathered in local surveys, they are not directly related to the natural environment people live in [22]. This way, it could be possible to establish target norms, or behaviors that the new technology intends to promote.

The ontology model, part of the SAM, maps the vocabulary of the problem domain and graphically represents the nature of reality by ontological (and temporal) dependencies between affordances and agents [30][10]. Agents (graphically represented by ellipses) are affordances that can take responsibility for their actions and actions of others [27]. Both agents and affordances can have attributes called determiners. Figure 1 illustrates the ontology model of this study considering firstly the current state (in gray), then the possible future reality by means of a new technology (in white).

In the current reality, a particular Society is the root agent of the model, which means it affords, i.e. determines how we conceptualize and perceive the "natural environment" and the natural resources that the agent Energy generator uses in the generation process, such as water from dams. This energy is then distributed to existing buildings by the Energy distributor, which means that the energy distribution ontologically depends on the natural resources and energy produced by the Energy generator agent. The affordance energy consumption, which is represented in kWh, depends on the distribution and on the agent person (who consumes it) to exist. The Energy metering affordance depends on the consumption and distribution to measure and to show the consumption through its display. Reading this display to generate the

bill is an Energy distributor responsibility, while the bill payment is a householder's responsibility. The energy consumption impacts the natural environment via energy generation. However, the person who consumes energy in households, schools, etc., is not aware of this connection.

In the new scenario, represented in white, a public display (the SEETree) appears as an affordance of buildings. It consists on a public installation located in a building that is attended by groups of people. By means of the public display, people may establish an agreement and configure collective saving actions by setting the level of consumption reduction, the number of days to achieve this challenge, and identifying those buildings with smart metering that will join the challenge. The collective action is intended to create an awareness of the impact of the energy consumption in the natural environment, possibly motivating people to more conscious energy consumption; the target people are both: who are part of groups, and those motivated by them.

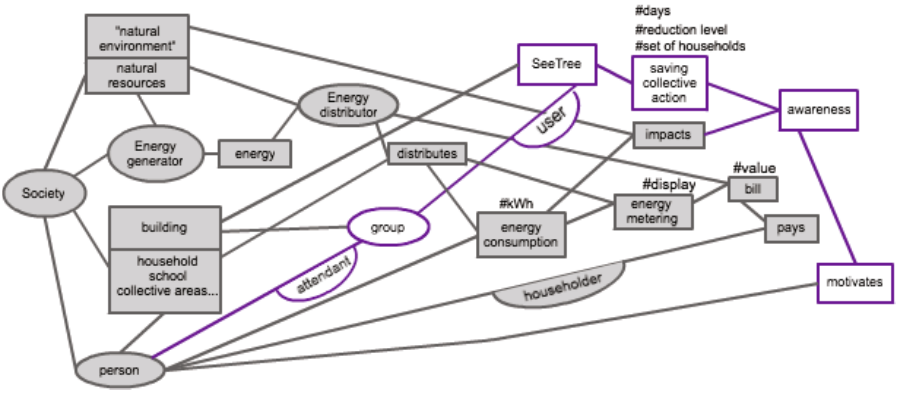


Fig. 1. The Ontology Chart of the relation with energy and the Social EFT proposal

The ontology chart (OC) in Fig. 1 differs in some epistemological aspects from a conventional OC [10][17] or "Stamper's OC"[34]. We thus call the type of OC presented in Fig. 1 *OC for Intended Reality* (OC4IR). Table 1 lists the main differences between the two types. In short, an OC4IR extends an OC with a part that models a possible future reality that is intended by the designers and possibly other existing or new stakeholders. Since it might introduce new affordances, agents, and ontological dependencies, it is conceivable that there are additional stakeholders in this future reality. The OC4IR is an intermediary document: once the innovation is being designed and disseminated, the OC4IR serves as the blueprint for the next OC that then represents the "new" actual reality. Consequently, it can be expected that the "new" OC deviates in some aspects from the OC4IR, as new stakeholders create a shared understanding of the new reality. For instance, if, for different reasons, awareness of the impact of energy consumption is still not afforded as expected, this will be reflected in the new OC. Furthermore, even parts of the "old" OC might

change, e.g., if new stakeholders with new insights are introduced that result in a different shared understanding. The most notable difference between an OC and an OC4IR lies in the "affordances" of the intended future reality.

Table 1. Comparing Stamper's OC with the OC for Intended Reality

	Stamper's OC	OC4IR
What is modeled?	An ontology in the philosophical sense, i.e. a shared representation of actual reality.	Apart from actual reality, a possible future reality as intended by designers.
Who subscribes to the model?	All involved stakeholders.	All stakeholders subscribe to the part that models actual reality, at least designers subscribe to the possible future reality.
Purpose of the model?	To document a shared model of actual reality (the Information System) and to eventually create a computational representation of the technical Information System.	Document the designers' (and possibly other stakeholders') understanding of the innovation task. Inform design: define new stakeholders, affordances, norms, and assumptions about the future, leading to requirements.
Agent	As defined in [10, 30].	As defined in [10, 30]; new agents become stakeholders.
Affordance	As defined in [10, 30].	Affordances of actual reality (A_{act}) as defined in [10, 30]; "affordances" of possible reality (A_{pos}) still need to be established and made perceivable → core challenge of design intention.
Ontological dependency	As defined in [10, 30].	As defined in [10, 30].
Norm	As defined in [10, 30]; prevalence of behavioral norms, expressed in deontic logic.	Additionally: other types/categories of norms regarding intentions, expectations or motivations.

4 Norms

An organization can be seen as a system of social norms, which have the function of determining patterns of behavior, and whether they are legal or acceptable within the social context, defining a culture. Norms are developed as collective affordances through practical experiences of people (agents) in the society, influencing how people perceive the world, make judgments and possibly guiding their behavior according to a subjective evaluation of a situation. Norms can be manipulated, applied and disregarded accordingly [35].

Behavioral norms, for example, are described according to the structure of an explicit rule

for a certain community and a certain purpose,
if x then A is (obliged/permitted/forbidden) to do y

where x is some perception of the situation, A is a responsible agent (a person or group of people) and y is the action. The agent A can only act in accordance with the norms of the community for the given purpose if he/she/it has the information necessary for perceiving the situation and the power to communicate other information where the action calls for it [31].

Table 2 summarizes norms classification found in OS literature [17][34]:

Table 2. Types of norms [17][34]

Criteria	Types of norms	Description
Formality	Informal	Norms that are known by people who can live according to them without their being able to express them in writing.
	Formal	Norms that can be performed by people following explicit written norms or *rules* which they can be trained to perform in a rather mechanical way.
	Technical	Norms that are handled in this way or are so exactly specified as to be capable of automation.
Aspects of the behavior they govern	Perceptual	How people receive signals from the environment via their senses through media such as light, sound and taste.
	Cognitive	Enable one to incorporate the beliefs and knowledge of a culture, to interpret what is perceived, and to gain an understanding based on existing knowledge.
	Evaluative	Help explain why people have certain beliefs, values, and objectives. Vary between cultures.
	Behavioral	Govern people within regular pattern.
	Denotative	Direct the choices of signs for signifying; such choices are culture-dependent, such as certain colors to signify happiness.
The effects of their execution	Standing orders	Result in a change in the physical world, because they are commands to perform actions, expressed as one may, may not, must or must not do something.
	Status	Define social structure and legal relations, designating liability, right or no-right over certain events or actions.
	Powers of intervention	Invoke or inhibit the use of existing standing orders.
	Legislative powers	Change other norms.
Types of objects that they are applied to	Substantive	Core business function and operation.
	Communi-cation	Patterns, structures and procedures of communication within an organization.
	Control	Introduce sanctions and rewards, acting as a mechanism to reinforce that everyone does what he or she is supposed to do, as prescribed by other two types of norms.

According to the Semantic Analysis and the OC4IR, in the next session norms from the actual reality and intended reality are described.

4.1 Identified Norms and Behavior

In preliminary work [21], a set of norms was identified based on interviews with energy company employees, most of them regarding commercial relationship between customers and the energy company, i.e., how people prioritize energy bill payment among other monthly bills, or reasons to adopt an illegal connection instead of paying for the energy regularly. Part of this set of norms is considered here because they express values and, consequently, how people relate to energy.

Beyond this qualitative analysis, two surveys done as part of Smart Grid deployment projects also subsidize norms identification. A total of 280 householders members of the projects answered a questionnaire of about 140 questions in the Brazilian state of Minas Gerais, which is partially urban and mostly rural; other 165 householders participated of the survey in the city of Rio de Janeiro, most of them from low-income areas. Interviews were conducted face-to-face by a research institute in 2011 in Minas Gerais and in 2012 in Rio de Janeiro. Results of the survey previously analyzed [22] and related to social life, environmental concerns, motivation, and how they relate to energy and to each other in daily life were selected to set norms.

Table 3 describes the norms and presents their classification from the Table 1, firstly in terms of formality informal, formal and technical levels, then according to the other types of norms. Entities represented in the first part are related to the traditional OC (actual reality), followed by those related to the OC4IR.

4.2 Discussion

From the collection of norms, we observe that the actual scenario is mostly constituted by formal and technical norms, while the possible future reality is dominated by informal norms, especially those related to the social psychological view. This is aligned to the intended aim of the new technology design of promoting behavior change. As for the aspects of behavior they govern, we notice more evaluative norms in the new set, suggesting more involvement of people in the joint action and collective achievement. Regarding the types of objects that the new set of norms are applied to, there is predominance of communication over control or substantive norms, as the whole scenario inherits the substantive norms of the OC. The norms description also indicates some elements to drive the design (underlined), such as green area, flooding, natural environment, trash disposal, etc.

Table 3. Norms description according to the actual reality and intended reality

Entity	Norm description	I/F/T	Type of norm
Actual reality norms (OC)			
Energy generator	ALWAYS when the energy consumption increases, THEN energy generator MUST produce more energy to supply the demand	T	Standing order, Substantive
	ALWAYS when the energy consumption is predicted to reach the limit, THEN the government HAVE TO plan new power plants construction to generate more energy	T F	Power of intervention
	SOMETIMES, when the demand for energy is close to reach the limit, THEN the Energy generator MUST activate thermoelectric power plants	F	Standing order. Power of intervention
Energy distributor	ALWAYS when a person wants to have energy in a building, THEN the Energy distributor MUST install an energy metering	T F	Status
	WHENEVER when the month is about to end, THEN the Energy distributor HAVE TO read the consumption registered by the meter	F	Substantive
Impact	MOST OF TIMES, when new power plants must be built, THEN new green areas HAVE TO be flooded to create new dams	T	Status
House-holder	MONTHLY, when the Energy distributor sends the bill, THEN the householder HAVE TO pay for the energy consumed	F	Status
	MOST OF TIMES, when the householder pays the bill, THEN he/she SHOULD be aware of the amount of energy they have consumed (they are not [21])	I	Cognitive
Person	MOST OF TIMES, when in contact with people who do not pay for energy THEN consumers MAY believe that they should not pay for energy too [21]	I	Denotative
	MOST OF TIMES, when having incomplete education level THEN a person may not make sense of concepts associated with the energy consumption and its bill (kWh), %, charts [21]	I	Cognitive
	FREQUENTLY, when having a household budget increase THEN low-privileged consumers CAN buy more new domestic electronic devices [21]	F	Evaluative

Intended reality norms (OC4IR)			
Public display	ALWAYS, when motivated to consume energy consciously THEN users MUST receive objective information from the public display about individual consumption that helps them to act accordingly	F	Behavioral, Communi-cation
	ALWAYS, when participating of a saving collective action THEN groups MUST see partial results of the action in a motivating way considering participants household measures	T	Control, Percepual
Saving collective action	ALWAYS when involved with a collective saving energy challenge THEN users MUST understand the importance of individual contribution to collective achievements	I	Cognitive
	ALWAYS when establishing a collective action THEN users MUST discuss and agree about each user and his/her family commitment	I	Evaluative, Standing order
	ALWAYS when establishing a collective challenge THEN group MUST determine a period to achieve it	T	Behavioral, Standing order
Awareness	MOST OF TIMES when using electricity in the household THEN users MAY think about natural resources consumptions	I	Evaluative Cognitive
	FREQUENTLY when using electricity in the households THEN users SHOULD discuss about energy consumption and the natural environment	I	Cognitive, Communica tion
	MOST OF TIMES when interested in preserving the natural environment THEN person SHOULD consider other natural resources, not restricting it to recycling/trash disposal [21]	I	Denotative
	MOST OF TIMES when wasting energy THEN individuals SHOULD relate it to environmental consequences	I	Evaluative Perceptual
	FREQUENTLY when interested in consciously consuming energy THEN families SHOULD check and discuss about the current consumption [21]	I	Behavioral Communi-cation
Motivates	ALWAYS when involved with a collective saving collective action THEN users MAY motivate other people in the society to consume more consciously [21]	I	Evaluative
	ALWAYS when involved with a collective saving energy challenge THEN users SHOULD engage their family in the commitment [21]	I	Behavioral
	FREQUENTLY when involved with a saving commitment THEN users MAY want to share their achievements with people they know [21]	I	Evaluative, Communi-cation

5 Conclusion

The relation between energy consumption and its environmental impact is weak or even almost inexistent in some sociocultural scenarios. This paper presented the concept of Social Eco-feedback Technology, drawing on the social affordance notion and on the analysis of data regarding the connection of energy consumption and the natural environment in Brazil. This approach to energy feedback differs from those in

the state of the art firstly by dealing with low-income areas in a developing scenario, and then by proposing a collective interactive solution for a public area.

The analysis and design process have been based on Organizational Semiotics artifacts, which helps to analyze information systems and the introduction of a technology within a social dimension. To consider the possible future influences by this technology, an adaptation of traditional artifacts of Organizational Semiotics was made. Further work involves designing a system that enables the informal and formal norms regarding collective saving interaction and then evaluating the technical solution proposal with regard to its potential for promoting the intended social behavior change.

References

1. Accenture. Understanding Consumer Preferences in Energy Efficiency, http://www.accenture.com/us-en/Pages/insight-understanding-consumer-preferences-energy-efficiency-utilities-summary.aspx
2. Banco Mundial. Consumo de Energia Eléctrica (kWh Per Capita), http://datos.bancomundial.org/indicador/EG.USE.ELEC.KH.PC.
3. Brazil. Ten Year Plan for Energy Expansion (PDE 2017), http://www.epe.gov.br/PDEE/Forms/EPEEstudo.aspx
4. Darby, S.: Smart metering: what potential for householder engagement? Building Research and Information 38(5), 442–457 (2010)
5. Dick, H., Eden, H., Fischer, G.: From consumers to owners: Using meta-design environments to motivate changes in energy consumption. In: Piccinno, A. (ed.) IS-EUD 2011. LNCS, vol. 6654, pp. 319–324. Springer, Heidelberg (2011)
6. Dourish, P.: HCI and environmental sustainability: the politics of design and the design of politics. In: Proceedings of the DIS 2010, pp. 1–10. ACM (2010)
7. European Commission. Energy Sources, Production Costs and Performance of Technologies for Power Generation, Heating and Transport, COM(2008)744. Brussels, Belgium (2008)
8. Froehlich, J., et al.: The design and evaluation of prototype eco-feedback displays for fixture-level water usage data. In: Proceedings of CHI 2012, pp. 2367–2376 (2012)
9. Froehlich, J., Findlater, L., Landay, J.: The design of eco-feedback technology. In: Proceedings of CHI 2010, pp. 1999–2008. ACM (2010)
10. Gazendam, M., Liu, K.: The evolution of organisational semiotics: A brief review of the contribution of Ronald Stamper. Studies in Organisational Semiotics. Kluwer Academic Publishers, Dordrecht (2005)
11. Gibson, J.J.: The Ecological Approach to Visual Perception. Houghton Mifflin, Boston (1979)
12. Goldstein, J.N., Cialdini, R.B., Griskevicius, V.: A Room with a Viewpoint: Using Social Norms to Motivate Environ. Conservation in Hotels. J. of Consumer Research (October 2008)
13. Hall, E.T.: The Silent Language. Anchor Books (1959)
14. Jelsma, J.: Designing 'Moralized' products: theory and practice. In: Verbeek, P.-P., Slob, A. (eds.) User Behaviour and Technology Development, pp. 221–231. Springer (2006)
15. Johnson, P.H., et al.: Beyond kWh: Myths and fixes for energy competition game design. In: Proc. of the Meaningful Play Conference (2012)

16. Kolkman, M.: Problem Articulation Methodology: PhD thesis, University of Twente, Enschede (1993) ISBN 90-9005972-5
17. Liu, K.: Semiotics in Information Systems Engineering. Cambridge University Press, Cambridge (2000)
18. Norman, D.: Affordance, conventions, and design. Interactions 6(3), 38–43 (1999)
19. Norman, D.: Simplicity is not the answer. Interactions 15(5), 45–46 (2008)
20. Norman, D.: The Design of Everyday Things. Doubleday, New York (1988)
21. Piccolo, L.S.G., Baranauskas, M.C.C.: Motivational Aspects in Energy Feedback Systems Design. In: Proc. of IHC+CLIHC, pp. 313–328. ACM (2010)
22. Piccolo, L.S.G., Baranauskas, M.C.C.: Climbing the Ladder with Energy: informing the design of Eco-Feedback Technology with a Social Approach. In: Proc. of the 14th Intl. Conf. on Informatics and Semiotics in Organisations (ICISO 2013). Springer (to be published 2013)
23. Pierce, J., et al.: Some consideration on the (in)effectiveness of residential energy feedback systems. In: Proceedings of the DIS 2010, pp. 244–247. ACM (2010)
24. Pierce, J., Paulos, E.: Beyond energy monitors: interaction, energy, and emerging energy systems. In: Proceedings of CHI 2012, pp. 665–674. ACM (2012)
25. Pierce, J., Schiano, D.J., Paulos, E.: Home, habits, and energy: examining domestic interactions and energy consumption. In: Proc. of CHI 2010, pp. 1985–1994. ACM (2010)
26. Rokeach, M.: Long-term Value Change Initiated by Computer Feedback. In: Rokeach, M. (ed.) Understanding Human Values: Individual and Societal, pp. 210–225. The Free Press (1979)
27. Salter, A., Liu, K.: Using Semantic Analysis and Norm Analysis to Model Organisations. In: Proceedings of ICEIS (2002)
28. Sanders, K.R., Atwood, L.E.: Value Change Initiated by the Mass Media. In: Rokeach, M. (ed.) Understanding Human Values: Individual and Societal, pp. 226–240. The Free Press (1979)
29. Spagnolli, A., et al.: Eco-Feedback on the Go: Motivating Energy Awareness. Computer 44(5), 38–45 (2011)
30. Stamper, R.K.: Organisational semiotics: Informatics without the computer? In: Information, Organisation and Technology: Studies in Organisational Semiotics, pp. 115–171. Kluwer Academic Publishers, Boston (2001)
31. Stamper, R.K.: Social Norms in Requirements Analysis - an outline of MEASUR. In: Requirements Engineering, Technical and Social Aspects. Academic Press, N.Y (1994)
32. Stamper, R.K., Althaus, K., Backhouse, J.: MEASUR: Method for Eliciting, Analyzing and Specifying User Requirements. In: Computerized Assistance During the Information Systems Life Cycle. Elsevier Science Publishers, North-Holland (1988)
33. Stamper, R., et al.: Semiotic Methods for Enterprise Design and IT Applications. In: Proc. of the 7th Int. Workshop on Organisational Semiotics, pp. 190–213 (2004)
34. Stamper, R., Liu, K., et al.: Understanding the Roles of Signs and Norms in Organisations: a semiotic approach to information system design. Journal of Behaviour & Information Technology 19(1), 15–27 (2000)
35. Stamper, R.K.: Social Norms in Requirements Analysis – an outline of MEASUR. In: Jirotka, M., Goguen, J., Bickerton, M. (eds.) Requirements Engineering Technical and Social Aspects. Academic Press, New York (1993)
36. US Department of Energy. The Smart Grid: An introduction, http://www.oe.energy.gov/DocumentsandMedia/DOE_SG_Book_Single_Pages.pdf
37. WWF International. Living Planet Report WWF, Gland, Switzerland (2012)

Coupons as Monetary Incentives in Participatory Sensing

Andreas Albers[1], Ioannis Krontiris[1], Noboru Sonehara[2], and Isao Echizen[2]

[1] Goethe University Frankfurt, Chair of Mobile Business & Multilateral Security
Grueneburgplatz 1, 60323 Frankfurt, Germany
[2] National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
{andreas.albers,ioannis.krontiris}m-chair.net,
{sonehara,iechizen}nii.ac.jp

Abstract. Participation of people is the most important factor in providing high quality of service in mobile sensing applications. In this paper we study coupons as incentives in order to stimulate users participation, especially in applications that rely on real-time data. We argue that coupons do not only function as incentives to increase user participation, but they can also direct more people to the targeted sensing area, increasing the overall utility of data for service providers. In this paper we study coupons in combination with multi-attributive auctions, which gives the additional advantage to service providers of not having to determine the coupon value that users would expect in exchange for their data. Instead users have to compete with each other to win the auction, choosing coupons of lower values. Even though the combination of coupons with multi-attributive auctions is very attractive for participatory sensing, we also highlight some of the problems coupons have and especially those connected with user privacy.

1 Introduction

The wide adoption of mobile devices in combination with the spread of the Web 2.0 paradigm on the Web recently created the right conditions for a new scope of research, often referred to as mobile phone sensing [1] or participatory sensing [2]. It complements our previous efforts in wireless sensor networks and due to sensor-rich devices, geo-localised user-generated content can now be created any time and anywhere. Other sensors, besides geo-location chips, such as camera, gyroscope, light sensor or accelerometer started to become more and more prevalent in mobile devices, enabling new large-scale practices.

An interesting class of applications that emerged from these advancements shares the vision of a sensor data-sharing infrastructure, where people and their mobile devices provide their collected data streams in accessible ways to third parties interested in integrating and remixing the data for a specific purpose or campaign. A popular example is a noise mapping application, which generates a collective noise map by aggregating measurements collected from the mobile

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 226–237, 2013.

phones of volunteers [3,4]. It can raise citizen awareness of noise pollution levels, and aid in the development of mitigation strategies to cope with the adverse effects. In similar applications, people monitor air pollution [5] or road and traffic conditions [6].

A key factor for the success of such applications lies in the people's participation in data sensing activities. Even though the ubiquity of mobile devices makes mass participation feasible, it remains questionable how the general public can be motivated to participate, especially since efforts that call people to freely volunteer have only been moderately successful so far.

From the viewpoint of service providers that collect sensing data, monetary incentives could increase user participation for the service. It could help service providers attract a large number of participants and thereby increase not only the collected amount of data, but also its quality, in order to offer a higher quality of service back to the users. Monetary incentives can be of the form of micro-payments, but they can become even more attractive in the form of coupons, which are nowadays easily transferable in electronic form on consumers' mobile phones and offer additional advantages from the perspective of service providers.

The challenge then becomes how to determine the price that users expect to receive for their effort of collecting data. This price may depend on the individual preferences of the users and how they perceive the cost of participating in the process. Sensing takes time, interrupts other activities, consumes additional battery power and data traffic bandwidth, and most importantly may require users to give up some of their privacy [7]. In general, each user has her own valuation and thereby a minimum price that he expects to receive, depending on how much effort and personal information he has to give out, in order to collect and provide sensing data for a service provider. This valuation naturally differs among individuals and based on the context or situation they are currently in.

In this paper we argue that an attractive solution to the above problems is the use of multi-attributive auctions, because service providers avoid determining the price, but they rather delegate it to the users, who are biding and competing with each other, pushing the prices down. In our solution however, we use coupons instead of micro-payments, which offer additional advantages. One of the main advantages is that coupons can incentivise people to move towards the targeted area, where we want to collect data from, and at the end increase the flow of data to the service provider. However, coupons also introduce additional difficulties and have an impact on how the auction mechanism should be organised. In particular, in this paper we study how coupons can be integrated into multi-attributive auctions and how the utility of the data at the service provider is affected.

The rest of the paper is structured as follows. In Section 2, we argue why multi-attributive auctions are the appropriate pricing mechanism for participatory sensing, we argue about the impact of using coupons instead of micro-payments and we describe analytically the steps of the auction mechanism with coupons. In Section 3, we discuss the parameters that affect the valuation of coupons by users and what additional context parameters can affect the overall utility of

the submitted data by the service provider. Section 4 presents our experiments results and Section 5 highlights the privacy concerns associated with coupons. Finally, in Section 6 concludes the paper.

2 Market-Based Acquisition of Sensing Data

Reddy et al. [2] demonstrated with user trials that a set of micro-payments can be effective in encouraging participation in mobile sensing. They experimented with fixed prices and concluded that having a fair micro-payment (20 or 50 cents per valid submission) with an achievable maximum pay-out is a good strategy for a balanced participation. In this regard, micro-payments constitute very strong means to influence users in favour of a service provider.

However, service providers offering fixed prices for the acquisition of sensing data cannot address the diversity of the resource "sensing data". As discussed above, the price that users expect to receive for their efforts is different among individuals and depends on many parameters. So, a better alternative for a pricing scheme would be to use auctions that could help to reveal the expected, unknown price of a user for her contribution.

Lee and Hoh [8] proposed recently the use of Reverse Auction based Dynamic Price (RADP) as a pricing mechanism in participatory sensing. Users bid for selling their sensing data, the service provider selects a predefined number of users with the lowest bid and the winners receive their bid prices for the data. Reverse auctions lead to the decrease of prices, since sellers (mobile users) compete with each other and continuously decrease their bids to increase their chances to win [9]. The winners of the auction round are able to sell the sensing data to the service provider at the price of their bid.

However, this is a sub-optimal solution, because a user's auction bid would consist only of the price a user expects to receive for her sensing data and does not allow to negotiate on the data quality as well. Here we need to take under consideration that sensing data are not all of same quality. In order for a service provider to provide a better service, collecting data of high quality is important[10]. For example, better location precision or higher sampling frequency are factors that are of interest to the service provider. Negotiating on the price alone does not provide any control over the quality of the data and therefore does not help service providers meet their application and quality of service requirements.

In [11], we showed that multi-attributive auctions (MAAs) are an attractive solution in the paradigm of participatory sensing, since they can help service providers select the most suitable sensing data, but also provide users with the incentive to improve on them. This mechanism provides attractive benefits for participatory sensing and in particular:

(i) Service providers can avoid the challenge of determining the price a user expects to receive for her sensing data.
(ii) Service providers can influence the data quality of the sensing data through the auction process, in order to meet their application requirements.

(iii) Users are able to evaluate and improve on their data quality during the auction process, in order to archive a higher bid.

Multi-attributive auctions have been typically applied so far in business-to-business procurement cases, but they have never been applied in cases where simple users are involved in the bidding process. By doing so, we have a completely new setting of roles, where simple users are the sellers and the service provider is the buyer of a virtual good "sensing data". This allows us to address the complex nature of incentives, compared to the limitations of "classical" pricing mechanisms and achieve at the end higher utility.

2.1 Replacing Micro-payments with Coupons

Monetary incentives used to motivate users in order to participate in activities for a service provider typically consist of money in the form of micro-payments. Another form of payments in this context is coupons, which are handed out as part of loyalty programs. This kind of coupons allows users to get discounts on selected products at participating retail stores (e.g. buy two - pay one). In recent years, such coupons have become more and more digital, which makes them easy to transfer electronically to a mobile wallet or to a loyalty eCard of a user (i.e. similar to digital money transfers). Nevertheless, coupons differ from real money in certain aspects, which in our context can provide additional benefits to service providers wanting to reward users in exchange for sensing data.

Although, there is great variety of different types of coupons available in practice, they all share three common characteristics: coupons can only be redeemed for specific products (e.g. lunch meal), at a specific retailer (e.g. Italian restaurant) and in a specified context (e.g. only in restaurants in a certain city within the next two month). Service providers can leverage these characteristics to their advantage in order to increase the utility of their collected sensing data, if they are also willing to address the complementary drawbacks.

One of the advantages is the possibility to make coupons location-dependent and motivate users to collect required data from a specific geographic region, by handing out coupons, which are only valid in a certain part of a city (around this area) and have at the same time a short expiration date. This can motivate users who are already in the area to sense data or most importantly make users move to that targeted area from other points. Existing research has already demonstrated empirical evidence for this through user experiments on the factors leading individuals to redeem coupons [12]. The findings suggest that the main drivers for this are the distance to the shop, the amount of discount, the time of day and the matching of coupon with user profile.

An addition benefit of using coupons is that service providers act as distributors/promoters of coupons and may receive a monetary compensation from retailers in exchange. This allows them to lower the total sensing data acquisition costs.

One should note here that even though the fact that a coupon can only be redeemed for a certain purpose is beneficial for retailers, it is at the same time a

drawback for the user. While money can be spent virtually everywhere, a coupon can only be used for certain purpose (product, brand, etc.), which a user may or may not value at all. Thus, coupons are always less worth to users than money.

Besides this obvious challenge for service providers to leverage the advantages of coupons while taking care of the drawbacks, coupon characteristics also make it more challenging to integrate them in multi-attributive auction mechanisms. As we said above, we employ MAAs, in order to determine the price that a user demands for her sensed data. If real money is used in the auction, the user can simply specify the price at which he is willing to sell. Coupons, on the other hand, have two monetary values. The first value (internal value) is determined by the objective amount of money a user can save by redeeming the coupon. The second value (external value) is determined by the personal valuation of the user for such a coupon. That is, the value of a coupon depends on the value a user sees in a product, for which the coupon offers a discount. Consequently, the challenge for service providers is to consider both values of a coupon in the multi-attributive auction mechanism.

In the next section we suggest a solution of how coupons can be integrated in multi-attributive auctions and we show analytically the steps of one auction round applied in the context of participatory sensing applications

2.2 Auction Process

We assume that a service provider wants to provide real-time services, such as monitoring noise, temperature or CO_2 levels in an urban environment. In order to address the shortcomings of covering such a wide geographic area with dedicated sensor nodes, the provider takes the approach to collect the data from mobile phones of individuals. The idea, then, is to orchestrate the computing, communication, and sensing capabilities of a population of mobile phones, which happen to be at the area of interest or a nearby area, in order to enable large-scale sensing purely through software running on this existing hardware base. Each mobile phone transmits the sensing data from the environment to a central server through the mobile network. Then, the service provider aggregates the measurements from all users and processes them to deliver the service.

Fig. 1 depicts the three more generic phases of the process, but in what follows we explain in details the specific steps involved.

Step 1: In order for the service to be real-time, the provider defines successive short time periods, within which sensing data are to be collected by people from a given geographic area. This is expressed by the service provider in the form of a task, which is submitted to a Task Server. The task contains two things: the *acceptance conditions* and a *utility function* $S(x)$. The acceptance conditions limit which mobile nodes may execute the task, e.g. it defines the required sensors, termination conditions, etc. The utility function incorporates the provider's budget and several parameters that affect the required data quality for the service provider. Basically, it allows the service provider (buyer) to communicate its quality requirements for a good, so that the mobile user (seller) is able to address these requirements alongside the submission of her monetary bid.

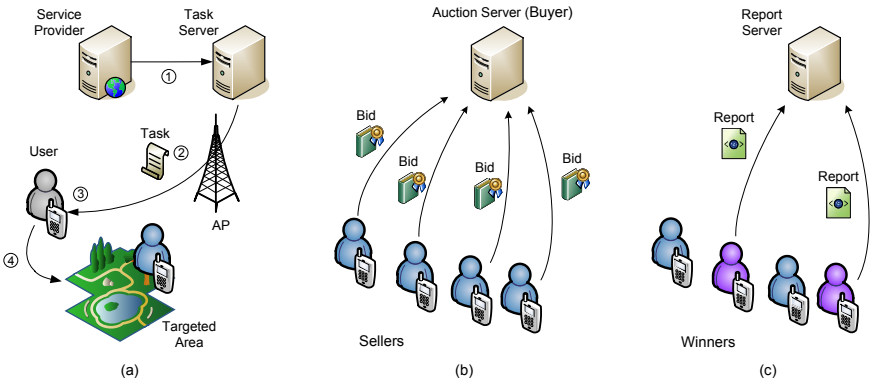


Fig. 1. The three phases of applying multi-attributive auctions for the acquisition of sensing data

Step 2: The mobile devices of users, registered in the service, periodically check on the Task Server and choose tasks to run based on the acceptance conditions of these tasks. Additionally, some local conditions, defined by the users, can also be applied. Such a condition could be, for example, not to accept sensing tasks, if the remaining battery level is below a threshold.

Step 3: If the mobile phone accepts a task, users define their initial price. As we discussed in Section 2.1, since service providers are giving out coupons instead of real money, specifying an actual price as in regular auctions does not work anymore. Instead, the user is presented with a set of coupons to choose from. Thereby it is essential for a service provider to select carefully, which coupons should be in such a predefined set. The challenge is to match the interests of the service provider and the user. The interest of the service provider is to motivate users to sense data in a certain geographic region and (if applicable) at the certain time. The user on the other hand is interested in receiving coupons that are close to her current location and match her personal product preferences. Consequently, if a user accepts a sensing task, service providers need to select coupons based on a user's current location and personal preferences profile.

However, coupons are only the monetary part of the user's bid (i.e. the internal value of a coupon that the user chooses). Multi-attributive auctions integrate additional attributes of a good into the auction bid, besides the price. So, each bid comprises a monetary bid as well as a multiple quality dimension. In particular, a bid is represented as n-dimensional vector Q of both monetary and non-monetary relevant attributes.

In the case of an additive utility function $S(x_j)$, the bid of the user (seller) can be expressed as $x = (x_1, \ldots, x_n)$. The service provider (buyer) evaluates each relevant attribute x_j through the utility function. As a result, the function $S : Q \rightarrow IR$ translates the value of each attribute into a *utility score*. At the end, the overall utility $S(x)$ for bid x constitutes the sum of all individual utility scores resulting from each attribute. If applicable, the individual utility scores

can be weighted, with the weights $w_1, \ldots, w_n$ summing up to the value of one. The overall utility of bid is given by Equation 1.

$$S(x) = \sum_{i=1}^{n} w_i S(x_i) \text{ and } \sum_{i=1}^{n} w_i = 1 \,. \tag{1}$$

So, in this step the utility function $S(x)$ is evaluated, in order to produce the utility score. This operation is locally at the device and does not involve any communication. The mobile device already has all the information needed to compute $S(x)$. In Section 3, we discuss several concrete parameters that could function as attributes of the utility function.

Step 4: If not satisfied by the utility score, users can improve it, by acting on the data quality factors (i.e. on the attributes x_i). For example, a user could choose a coupon with smaller value, or blur the location information (for location privacy protection), or move closer to the preferred sensing area of the service provider (as shown in Fig. 1(a)) and provide more accurate data.

Step 5: The utility score is recalculated locally at the user's device and once the user is satisfied with the value, he submits her bid to the Auction Server for the current auction round, as shown in Fig. 1(b). Her bid is not the data itself, but rather the utility score, i.e., the characteristics of the sensing data. The auction round begins at the end of each reporting time period and takes place between the users who have submitted their bids in that round.

For m submitted bids, the buyer determines the winning bid to be the one with the maximum utility score, as given by Equation 2.

$$\max\{S(x_j)\}, \text{ where } 1 \leq j \leq m. \tag{2}$$

However, in most applications it would not be enough to have only one winner and therefore only one data report. In that case, Equation 2 can be generalised to select the bids with the n biggest utility scores and that consequently determines n winners for the auction round.

Step 6: After the auction round winners have been determined, they get notified by the service provider and they submit the actual sensing data from their devices to the Report Server. This is depicted in Figure 1(c). The rest of the users who lost this auction round do not submit their data, but their probability to win in the next round is increased, as we will see in the next section.

Step 7: Finally, the service provider transfers the coupons to the winning users. If the users actually redeem the coupon in the specified retail store, their action is fed back to the service provider as well as the retailer. For the service provider, this information is valuable in terms of refining the preference profile of the user in order to suggest relevant coupons for a user to choose from in the next auction round. In that way we keep the external coupon value at a high level, which would be necessary for the coupons to keep working as incentives. At the same time, for retailer this information provides a variety of insights as well (e.g. at which location and time has a user received the coupon, did he actually redeemed the coupon, etc.)

3 Assigning Attributes to Sensing Data

As we saw in the previous section, the service provider (buyer) defines a utility function $S(x)$ beforehand, including multiple attributes x_i that is important to him and on which users will compete during the auction. Each attribute is weighted differently, by assigning the corresponding weight w_i. Then, the question becomes: what are those attributes that the service provider can include in the utility function? The obvious one is the selling price requested by the users, which the service provider wants to keep low and within the available budget. Below we discuss a list of additional attributes. We use the Quality of Context framework of Buchholz et al. [13] to derive a selection of such attributes for quality of sensing data. It is not meant to be a complete list, but only give an indication on what those attributes could be.

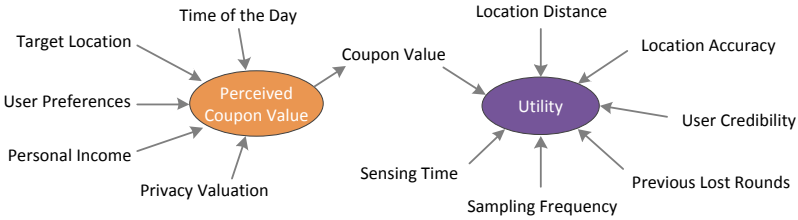


Fig. 2. Parameters affecting the utility function, $S(x)$

Coupon Value: The coupon value needs to be described from two perspectives. First, there is an internal value for each coupon. It denotes the actual costs of a retailer in the case a user actually redeems a coupon (e.g. 30 percent off). Second, there is an external value of a coupon, which reflects its personal valuation by a user. This valuation is determined by several factors:

(i) Time of the Day: Coupons need to be redeemable in short period of time for users. If this is not possible (e.g. due to certain opening hours of a retailer), this may decrease the value of a coupon for a user.
(ii) Target Location: The same as above applies to location. If the distance between the issuing retail shop and the user is too large, it takes correspondingly long time for the user to redeem this coupon. Consequently, it decreases the value of a coupon for that user.
(iii) Personal Income: A coupon allows a user to save a certain amount of money. The perceived value of this is determined by the personal income of that user: The higher the income, the lower the external value of the coupon.
(iv) User Preferences: Each user has personal preferences about certain products and he values them differently. Consequently, the external value of the coupon for a user is determined by how close the coupon manages to address her personal preferences.

(v) Privacy Valuation: In order to offer users relevant coupons, service providers have to know the current location and personal preferences of a user, which directly affects her privacy. Then, the perceived value of the coupon for that user will be affected by this trade-off and her overall privacy attitude.

Eventually, the external value determines the internal value of a coupon, since the users are given the opportunity to choose amongst a set of coupons. A user will always choose a coupon based on its external value (i.e. the personal value of a coupon for him), which determines the internal value and therefore the costs for a retailer/service provider. Nevertheless, during the auction the user is able to decrease her bid, in order to improve her winning position. That corresponds to choosing a coupon with smaller value out of the available coupons and resubmitting her bid, as described in Section 2.2.

Amount of submitted data: The sampling frequency and the overall time the sensors are measuring the environment affect the energy resources of the user's mobile phone. By increasing them, the user would devote more resources, but also increase the data quality and therefore the utility score.

Sensing location distance: Service providers typically prefer a specific location, at which sensing data is to be collected by mobile users. In order to decrease the sensing location distance, a mobile user could decide to physical move closer to the targeted location. This improves the quality of the sensing data for the service provider and at the same time the value of the user's bid in the auction.

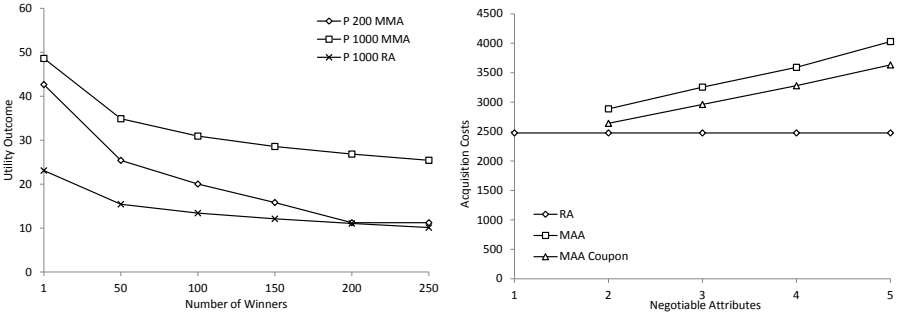
Location accuracy: Blurring location data is a common technique to provide location privacy [14]. This however could reduce considerably the worth of the data for the service provider. The mobile user could be offered the possibility to reduce the blurring factor and provide more useful data in exchange for a higher probability to win the biding process.

User credibility: Depending on the disclosed personal information of mobile users, a service provider is able to derive additional sensing data quality aspects. For instance, if an individual mobile user can be uniquely recognized by a service provider (e.g. based on a pseudonym) as opposed to anonymous users, the service provider has the opportunity to derive the creditability of the submitted sensing data, based on her reputation score and the available user transaction history.

Number of previously lost auction rounds: Based on the transaction history of a mobile user, it can be revealed how many times they failed in winning prior auction rounds. The higher this count, the higher the utility score a mobile user gets in their current auction round (the actual score depends on the individual utility function of a service provider). In this way, we motivate users to keep participating, even if they fail to win an auction from time to time. However, it needs to be transparent to the mobile users that they benefit from the fact they have previously failed to win an auction round.

4 Simulation Experiments

In our experiments, we compare the traditional Reverse Auction (RA) format with the multi-attributive auction format. The resource *sensing data* is described



(a) Utility dependence on the number of winning users.

(b) Acquisition costs as a function of the negotiable attributes with users.

Fig. 3. Simulation results

by the coupon value and a number of quality attributes, as described by Eq. 1. For the sake of simplification we have assumed equal weights for all attributes, unless stated otherwise. The values that these attributes receive would be defined by the actual users, but for the simulation experiment they were randomly generated based on uniform distribution, which is the most realistic assumption for multi-attributive auctions [9]. Based on the *Monte Carlo Simulation* model, 5000 auction rounds had been conducted for each simulation case below.

In the usual case in participatory sensing, we will require the data of several sources, in order to cover a geographical area efficiently. Fig. 3a shows how the utility for service providers changes, if more than one users can win the auction. In particular, the figure shows two cases for MAA (for 200 and 1000 overall participants). The figure also demonstrates how the utility score in MAAs compares with the simper case of Reverse Auction, where users negotiate only on one attribute (the price). In both cases the utility outcome drops as the number of winners increases. The reason for this lies in the fact that an increasing number of winning users also decreases the competition between the users. This ultimately leads to a lower utility. A service provider has to carefully balance between its desired data quality (due to the amount of sensing users) and the desired overall utility of the collected sensing data. Nevertheless, it is important to emphasise that the MAA case still provides a higher overall utility than the RA.

The second simulation case shows the change of the sensing data acquisition costs as a function of the available number of negotiable attributes by users. 50 users as sellers of sensing data and one service provider as buyer of the data were assumed. For simplicity, only one winner of the auction was assumed. The results are shown in Fig. 3b and they demonstrate that the acquisition costs for service providers increase, once we apply MAAs and the coupon value is no longer the only attribute to be negotiated. This is because users are no longer solely competing on the coupons and instead competition is now equally distributed across multiple attributes on sensing data. However, a service provider has the option to increase the weight for the attribute "coupon value" within the utility

function. This case is denoted as "MAA Coupon" in Fig. 3b. Thereby, the service provider is able to reach a trade-off between acquisition costs and utility outcome based on the requirements of the application and business model. This means, the higher the weight on the coupon value, the lower the acquisition costs and also the lower the overall utility outcome.

5 Discussion

The main advantage of replacing micro-payments with coupons for sensing data is to increase the attraction of users towards a desired sensing target area. Indeed, user surveys have shown that up to 68% of people find location-based coupons attractive [15]. At the end this would increase the utility of sensed data for service providers. In addition, coupons will generate additional revenues coming from coupon-issuing retailers, as we mentioned in Section 2.1.

However, these benefits come at the expense of user privacy. First of all, service providers have to be aware of the product preferences or immediate needs of a user, in order to be able to offer attractive coupons in exchange for sensing data. This kind of information typically has to be provided by retailers and enables access for the service provider to all user preference information from related retailers to which a user has subscribed.

In addition, once users redeem a coupon while buying a product, they disclose their product preferences as well as their current location (i.e. location of the retailer). This privacy problem does not only link to the retailers but also to the sensing platform provider. This is because the retailer will eventually forward this information back to the service provider, in order to improve its future ability to offer attractive coupons to users (closed-looped marketing).

In order to develop a viable business model for the service provider, the outlined privacy issues need to be addressed by further research that allows to balance between stakeholder benefits (i.e. service providers, retailers, and users) and privacy protection. As there is currently no good solution for this, it is also important to equip users with the ability (e.g. transparency and control) to make informed decisions on the trade-off between monetary incentives and privacy.

6 Conclusions

The combination of coupons with multi-attributive auctions can create a powerful solution for service providers that want to offer monetary incentives to mobile users in the contact of participatory sensing applications. It enables service providers to negotiate with users for a higher quality of sensing data. Especially the use of location-based coupons can incentivise more people to move in the targeted sensing data increasing the utility of data even further. Furthermore, service providers have additional economic benefits, since users compete with each other and drop the prices they ask for their data. Coupons bring additional revenue coming from the retailers issuing the coupons. However, coupons introduce also important privacy threats for the users that need to be addressed. User

trials are required here, in order to evaluate how users would behave in practice in their interaction with the MAA mechanism and how this behaviour is affected by their privacy concerns. Our future research is targeted in this direction.

References

1. Lane, N., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., Campbell, A.: A survey of mobile phone sensing. IEEE Communications Magazine 48(9), 140–150 (2010)
2. Reddy, S., Estrin, D., Hansen, M., Srivastava, M.: Examining micro-payments for participatory sensing data collections. In: Proceedings of the 12th ACM International Conference on Ubiquitous Computing (Ubicomp) (2010)
3. Maisonneuve, N., Stevens, M., Ochab, B.: Participatory noise pollution monitoring using mobile phones. Information Policy 15, 51–71 (2010)
4. Kanjo, E.: NoiseSPY: A real-time mobile phone platform for urban noise monitoring and mapping. Mobile Networks and Applications 15, 562–574 (2010)
5. Honicky, R., Brewer, E.A., Paulos, E., White, R.: N-smarts: networked suite of mobile atmospheric real-time sensors. In: Proceedings of the Second ACM SIGCOMM Workshop on Networked Systems for Developing Regions (NSDR 2008) (2008)
6. Mohan, P., Padmanabhan, V.N., Ramjee, R.: Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In: Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys 2008), pp. 323–336 (2008)
7. Krontiris, I., Freiling, F.C., Dimitriou, T.: Location privacy in urban sensing networks: Research challenges and directions. IEEE Wireless Communications 17(5), 30–35 (2010)
8. Lee, J.S., Hoh, B.: Sell your experiences: a market mechanism based incentive for participatory sensing. In: Proceedings of the 8th IEEE International Conference on Pervasive Computing and Communications (PerCom) (2010)
9. Bichler, M.: The Future of e-Markets: Multi-Dimensional Market Mechanisms. Cambridge University Press, New York (2001)
10. Boutsis, I., Kalogeraki, V.: Dynamic QoS-aware event sampling for community-based participatory sensing systems. In: Proceedings of the 6th ACM International Conference on Distributed Event-Based Systems (DEBS 2012), pp. 153–156 (2012)
11. Krontiris, I., Albers, A.: Monetary incentives in participatory sensing using multi-attributive auctions. International Journal of Parallel, Emergent and Distributed Systems 27(4) (2012)
12. Spiekermann, S., Klafft, M.: Elektronisches couponing: Wetter schön hohe couponeinlöseraten. Absatzwirtschaft 8(2), 4–9 (2006)
13. Buchholz, T., Küpper, A., Schiffers, M.: Quality of context: What it is and why we need it. In: Proceedings of the 10th International Workshop of the HP OpenView University Association (HPOVUA 2003), Geneva, Switzerland (2003)
14. Kapadia, A., Triandopoulos, N., Cornelius, C., Peebles, D., Kotz, D.: AnonySense: Opportunistic and privacy-preserving context collection. In: Proceedings of the 6th International Conference on Pervasive Computing (May 2008)
15. Acardo: Akzeptanz von mobile couponing. Technical report, Acardo / lb-lab (2011)

Foundational Ontologies for Semantic Integration in EAI: A Systematic Literature Review

Julio Cesar Nardi[1,2], Ricardo de Almeida Falbo[2], and João Paulo A. Almeida[2]

[1] Research Group in Applied Informatics, Informatics Department,
Federal Institute of Espírito Santo, Campus Colatina, Colatina, ES, Brazil
[2] Ontology & Conceptual Modeling Research Group (NEMO), Computer Science Department,
Federal University of Espírito Santo, Campus Goiabeiras, Vitória, ES, Brazil
julionardi@ifes.edu.br, {falbo,jpalmeida}@inf.ufes.br

Abstract. Despite (i) the recognized benefits of using ontologies in semantic EAI initiatives, (ii) the benefits of using foundational ontologies for promoting meaning negotiation and common understanding, and (iii) the importance of the semantic integration issue in EAI area, foundational ontologies have not yet become widely adopted in EAI initiatives for dealing with semantic conflicts. This has led us to investigate, through a systematic review of the literature, the adoption of foundational ontologies in EAI initiatives, with the purpose of understanding the current role of these ontologies in EAI and identifying gaps for future research, in which the potential benefits of such ontologies could be explored. We consider: (i) the role of foundational ontologies as part of the integration approach; (ii) the use of ontologies at development time and/or at run time; and (iii) the adoption of systematic approaches for semantic EAI.

Keywords: Enterprise Application Integration, Foundational Ontologies, Semantic Interoperability, Semantic Integration, Systematic Literature Review.

1 Introduction

The Enterprise Application Integration (EAI) research area regards the development and use of plans, methods, and tools to put together distinct information systems, by supporting their ability to exchange information and functionality to accomplish specific parts of a collaborative business process [1][2]. By dealing with distinct software applications, from both intra- and inter-organizational contexts, semantic conflicts arise. These conflicts are caused, among others, because the various heterogeneous applications do not share the same conceptualization [3].

A foundational ontology is a kind of (meta)ontology, independent of a particular problem or domain, that describes a set of real-world categories that can be used to talk about reality. It is constructed using the theories of Formal Ontology in philosophy [4][5]. The main purpose of a foundational ontology is to aid in negotiating meaning and facilitating common understanding [6]. Because of that, foundational ontologies have been acknowledged as an important means for dealing with semantic conflicts [5].

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 238–249, 2013.

However, despite the recognized benefits of using ontologies in semantic EAI initiatives [3], the benefits of using foundational ontologies for promoting meaning negotiation and common understanding [6], and the importance of the semantic integration issue in EAI area [3], foundational ontologies have not yet become widely adopted in EAI initiatives for dealing with semantic conflicts. This has led us to investigate the adoption of foundational ontologies in EAI, aiming at understanding the current role of these ontologies in this area and identifying gaps for future research, in which the potential benefits of foundational ontologies could be explored.

In this paper we present a Systematic Literature Review (SLR) [7], analyzing EAI initiatives that adopt foundational ontologies to address semantic integration. We have formulated three research questions (see section 3), which aim to investigate three main aspects of the surveyed studies: (i) the role of foundational ontologies as part of the integration approach; (ii) the use of ontologies at development time and/or at run time; and (iii) the adoption of systematic approaches for these EAI initiatives.

This paper is organized as follows. Section 2 presents the theoretical background and clarifies some terminological aspects. Section 3 describes the research method adopted and the review protocol. Section 4 presents the selection process and a brief description of the selected studies. Section 5 presents a synthesis of the data collected from these studies in the light of the research questions. In Section 6, we discuss important points identified during data analysis, which may be useful for the research agenda in EAI area. Finally, in Section 7, we present some final considerations.

2 Background

Ontologies have been acknowledged as an important means to address semantics in EAI [3]. An ontology is a formal representation of a common conceptualization of a universe of discourse [4]. There are different classifications of ontologies. Guarino [4] classifies ontologies into: (i) foundational (or top-level) ontologies, which describe very general concepts independently of a particular problem or domain, such as object, event, action etc., (ii) domain ontologies, which describe the conceptualization related to a generic domain (for instance, law or biology), (iii) task ontologies, which describe the conceptualization related to a generic task (such as diagnosis or planning), and (iv) application ontologies that describe concepts dependent on a particular domain and a task. Scherp et al. [8] extend this classification by admitting the so-called “core ontologies” with a level of generality between that of foundational and domain ontologies. In this sense, core ontologies provide a precise definition of structural knowledge in a specific field that spans across different application domains.

Although these kinds of ontology are important for clarifying concepts used along this work, we are especially interested in foundational ontologies. Examples of foundational ontologies include DOLCE [9], SUMO [10], YAMATO [11], and UFO [5].

In an ontology-driven Information System (IS), an ontology can be used at development time or at run time [4]. At development time, the knowledge represented by a domain/task ontology can be reused to support conceptual analysis and assure the ontological adequacy of the IS. Further, a foundational ontology may be used at development time to support requirements engineers in conceptual modeling [4]. In this case, fundamental ontological distinctions embodied in the foundational ontology

are used to improve the quality of conceptual models [5].When using ontologies at development time, we should focus on their representation adequacy, since these ontologies, said *reference ontologies*, are used in an off-line manner to assist humans in tasks of meaning negotiation and consensus establishment. On the other hand, once users have already agreed on a common conceptualization, specialized versions of the reference ontology can be created for run-time use. These versions are said *lightweight ontologies*, and sacrify representation adequacy to guarantee desirable computational properties [5].

The principles that guide ontology-driven IS development can also be applied in ontology-based EAI initiatives. In EAI initiatives, ontologies may be used at integration development time, in the sense that the ontology can be used by humans to negotiate meaning during the integration process, more specifically during an activity of conceptual mapping between applications. So, the two scenarios described by Guarino [4] hold: domain and task ontologies may be used by humans to negotiate meaning when mapping concepts and relations between different applications, as well as foundational ontologies may be used to capture fundamental ontological distinctions in order to categorize applications' concepts when mapping them. On the other hand, ontologies may also be used at integration run time, in the sense that an ontology may be implemented as a lightweight ontology, in a machine readable language, and used to link the applications at run time.

Finally, in the context of this review, it is necessary to clarify some aspects of terminology and scope. Firstly, we are interested in both intra- and inter-enterprise application integration. So, we use the term "enterprise application integration" to cover both categories, since most methods and patterns apply to both of them [2]. Secondly, although the terms "integration" and "interoperability" have been used to refer to different or interrelated concepts, they have also been used in an indistinct way. In this paper, we use the term "integration" in a broad sense, encompassing all which is called integration and interoperability. In order to cover a wide range of intended senses, we consider both "application integration" as well as "application interoperability" in the search string presented in Section 3.

3 Research Method and Review Protocol

The research method for this SLR was defined according to [7]. A systematic literature review is a form of secondary study that uses a well-defined method to identify, analyze and interpret the available evidences in a way that is unbiased and (to a degree) repeatable. A secondary study is a study that reviews primary studies related to specific research questions with the aim of integrating/synthesizing the evidences related to these research questions [7].

A SLR involves three phases [7]: Planning, Conducting and Reporting the review. Planning involves the pre-review activities, which the most important one is defining the review protocol. Conducting the review is concerned with searching and studies selection, and data extraction and synthesis. Reporting is the final phase and involves writing up the results and circulating them to potentially interested parties.

The main parts of the review protocol used in this work are described as follows.

Research Questions. This SLR aims to answer the following research questions:

RQ1. How have foundational ontologies been used as part of EAI approaches?

RQ2. Do the studies use the ontologies at development time, at run time or both?

RQ3. Do the studies follow a systematic approach for performing the integration project? (Do they adopt or propose a method or a process model defining activities, inputs, outputs, guidelines, etc.?)

Question RQ1 aims to investigate if the ontological distinctions of foundational ontologies have been used directly to integrate applications or indirectly for grounding other lower-level ontologies that are, then, used to integrate the application. From that, question RQ2 aims to look if foundational ontologies have been applied at development time and/or at run time; Finally, question RQ3 refers to if the studies that use foundational ontologies present systematic integration approaches for addressing semantic EAI.

Search Strategy. The search strategy was split in two: EAI-FO and FO-EAI strategies. Both focus on finding studies that use foundational ontologies in order to address semantic enterprise application integration. However, the EAI-FO strategy starts looking for semantic EAI initiatives and then focusing on those that use foundational ontologies, whereas the FO-EAI strategy starts looking for the use of foundational ontologies to solve integration problems and then focusing on those that address EAI. We adopted this dual search strategy because, depending on the focus given by the authors, it is difficult to establish a single set of terms that is able to capture the full scope of our review. So, the EAI-FO and FO-EAI strategies are complementary and aim to offer a wider review strategy. As a result, we define two search strings (that were applied to title, abstract and keywords of each study), as follows:

EAI-FO search string

("application integration" OR "application interoperability" OR "enterprise system integration" OR "enterprise system interoperability" OR "integration of information system" OR "interoperability of information system" OR "integration of application" OR "interoperability of application" OR "interoperability of enterprise application" OR "interoperability of enterprise system" OR "integration of enterprise application" OR "integration of enterprise system" OR "interoperability of business application" OR "interoperability of business system" OR "integration of business application" OR "integration of business system" OR "integration of heterogeneous system" OR "integration of heterogeneous application" OR "interoperability of heterogeneous system" OR "interoperability of heterogeneous application" OR "interoperability of information system" OR "integrated application" OR "interoperable application" OR "integrated enterprise system" OR "interoperable enterprise system" OR "information system integration" OR "information system interoperability" OR "enterprise system integration" OR "enterprise system interoperability" OR "business system integration" OR "business system interoperability") AND (semantic OR semantics OR semantically)

FO-EAI search string

("foundational ontology" OR "top-level ontology" OR "top level ontology" OR "upper-level ontology" OR "upper level ontology" OR "upper ontology") AND ("integration" OR "interoperability")

Selection Criteria. The selection criteria are organized in inclusion and exclusion criteria. There is a single inclusion criterion: (IC1) The study presents a semantic EAI

initiative that uses foundational ontology. There are four exclusion criteria: (EC1) The study is not written in English; (EC2) The study is an older version of another study already considered; (EC3) The study is not a primary study (including editorials and summaries of keynotes, workshops or tutorials); and (EC4) The study is just published as an abstract. Moreover, we considered the studies published until December 31th 2011, since we performed the data collection in the beginning of 2012.

Sources. The following sources were searched: IEEEXplore, ACM Digital Library, Springer Link, Thomson Reuters Web of Knowledge, Scopus, Science Direct, and Compendex.

Assessment. In order to avoid biases along the review process, we performed periodic meetings. The review process was conducted by one of the authors, and one of the other two performed activities of verification and validation over samples that represent about 35% of the studies. Possible biases were discussed in these meetings.

4 Selected Studies

An overview of the selection process is presented in Fig. 1. Using the EAI-FO strategy, 702 records were retrieved from the considered sources. After that, we eliminated duplicate records and applied the selection criteria, resulting in 5 studies. Using the FO-EAI strategy, 227 records were retrieved and 3 studies remained after eliminating duplicate records and applying the selection criteria. In the intersection of the two result sets, there was 1 study. So, in the end, the following 7 primary studies remained: [2][8][12][13][14][15][16]. Next, we present a brief description of each one.

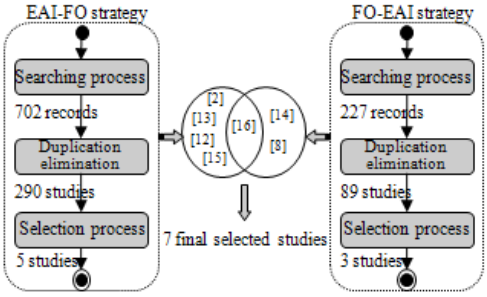


Fig. 1. The searching and selection processes

In [13], a use case scenario for integrating heterogeneous project management applications aims to test the ability of Conceptual Spaces (CSpaces) to handle integration problems. A CSpace is a knowledge container defined as a set of tuples, with a well-defined structure. In the use case, the PSL Ontology (Process Specification Language Ontology) [17] is stored in a CSpace that is shared by several applications. Also, each application has an individual CSpace that stores the extended version of the PSL Ontology together with application data.

The studies [2] and [12] were conducted in the context of the FUSION Project, and thus they share common principles. In [12], Bouras et al. present the ENterprise

Integration Ontology (ENIO), which is based on the alignment of DOLCE [9] and SUMO [10], with extensions to cover some dimensions of the EAI domain, called facets. There are three facets: data facet is taken as basis in transformations of services' inputs and outputs; functional facet defines the services' capabilities and provides classes for annotating services' operators with functional semantics; and process facet aims at providing a means for defining collaborative business process templates. In [2], Alazeib et al. describe two approaches for semantically-assisted design of collaborative business process (manual and semi-automatic). Both approaches rely on ENIO for solving structural heterogeneities of different business applications.

In [16], Treiblmayr et al. present an initiative for integrating a Geographic Information System (GIS) with a non-GIS application. These applications are integrated in terms of a service-oriented architecture, and the focus is on semantic interoperability. In order to map the conceptual schemes of the services provided by the different systems, ontological distinctions are analyzed at the light of DOLCE [9], or based on a semantic referencing system, also anchored in DOLCE.

In [14], Masuya et al. describe the development of the RIKEN integrated database of mammals. This database integrates heterogeneous mammal-related data stored in multiple individual databases of related research projects. To ensure the consistent classification of the content, they used a top-middle level ontology, YAMATO-GXO Lite. It is the lightened version of the middle-level ontology Genetics Ontology (GXO) to bridge the experimental genetics domain and the top-level ontology YAMATO [11].

Scherp et al. [8] discuss the use of three core ontologies in an integration scenario in the emergency response domain. The Event-Model-F is a core ontology representing different aspects of events. The Core Ontology on Multimedia (COMM) is designed for describing arbitrary digital media data. The Cross-Context Semantic Information Management Ontology (X-COSIMO) is designed for modeling semantic information management and communication. These three core ontologies are based on DOLCE, and follow a pattern-oriented design approach.

Finally, in [15], Paulheim and Probst present an ontology-based framework for integrating applications in the user interface level. For this, they use two separate domain-level ontologies: an ontology of the user interfaces and interactions domain, and a real world domain ontology. The latter is not an integral part of the framework and can be exchanged when integrating an application of a different domain. Any way, Paulheim and Probst argue that a "suitable real world domain ontology" is required, and in the scenario used to illustrate the proposed approach, the domain ontology is grounded in DOLCE. However, the ontology of the user interfaces and interactions domain is not grounded in any foundational ontology.

5 Data Synthesis

After selecting the primary studies, we analyzed each one in order to answer the research questions. Next, we perform the data synthesis regarding these questions.

RQ1. How have foundational ontologies been used as part of EAI approaches?
Analyzing the selected studies, we identified three different ways of using foundational ontologies in EAI initiatives: direct use of a foundational ontology (direct

approach), indirect use of a foundational ontology (indirect approach), and hybrid (both direct and indirect) use of a foundational ontology (hybrid approach).

In the *direct approach*, the ontological distinctions provided by foundational ontologies are directly used to solve semantic conflicts between the applications' concepts. This is the case of [13], where Martín-Recuerda uses the PSL Ontology as an interlingua for application integration.

In the *indirect approach*, a foundational ontology is used as basis for developing a core ontology or a domain ontology, which is, in turn, effectively used in the EAI initiative. This is the most common approach, and it is the case of [2][12][14][8][15], which are described below.

One of the indirect approaches is that employed in the context of the FUSION Project [2][12], in which the alignment of two foundational ontologies is part of the ENterprise Application Integration Ontology (ENIO). ENIO aims to provide an explicit specification of the EAI domain. It comprises, among others, a foundational ontology that is based on the alignment of DOLCE and SUMO. This foundational ontology is used for grounding an upper-level EAI ontology and the facets that cover three dimensions of the EAI domain: data, service and process. ENIO's facets present domain independent elements, which are based on widely adopted standards (and thus can be viewed as core ontologies), and domain dependent elements, which can be viewed as domain ontologies. The elements of each facet are used by EAI initiatives.

Another indirect approach is presented in [14]. Masuya et al. use YAMATO-GXO Lite, which aims to provide an explicit specification of the field of experimental genetics. This specification is used to integrate heterogeneous databases. As a result, 41 classes conveying the key information from each database were classified under 15 upper classes, which were used to classify the content of the integrated databases.

Scherp et al. [8] use and combine three core ontologies (Event-Model-F, COMM and X-COSIMO) to model the knowledge in an emergency response scenario. These core ontologies are aligned with DOLCE+DnS Ultralight2 (DUL), a lightweight version of DOLCE. Considering that the concepts of the core ontologies are used to integrate heterogeneous systems, this study is also classified as using indirect approach.

Finally, to facilitate ontology-based application integration on the user interface level, Paulheim and Probst [15] follow a plugin-based approach, in which each application to be integrated is encapsulated in a plugin. Application developers should formally describe their applications in application ontologies, which have to be based on domain-level ontologies as common ground. Also, the authors present an integration scenario in which a domain ontology is grounded by a foundational one (DOLCE), characterizing the indirect approach.

In the *hybrid approach*, a foundational ontology is used both directly and indirectly to aid solving semantic integration conflicts. This is the case of [16]. Treiblmayr et al. adopt an approach for concept mapping where concepts of each application are described based on both: (i) primitive concepts from a common shared vocabulary given by DOLCE (direct approach); and (ii) what they call a "semantic reference system" (indirect approach). In this context, the semantic reference system can be seen as a core ontology of the geographic information field, and its concepts are also grounded in DOLCE. Thus, the applications' concepts are anchored both in concepts from DOLCE and from the semantic reference system, characterizing a hybrid approach.

RQ2. Do the studies use ontologies at development time, at run time or both?
All the analyzed studies, in some level and stage of the integration process, consider a conceptual perspective when activities related to analysis and conceptual mapping are performed. Thus, all of them are classified as using ontologies at development time, mainly by integration engineers to negotiate the meaning of concepts and relations involved in the applications being integrated. Moreover, most of them also use ontologies at run time, implementing lightweight versions of the ontologies in a variety of machine-readable languages, such as KIF, RDF, OWL and F-Logics.

In [13], mapping and transformation rules between the applications' concepts and PSL are defined at development time and used to link CSpaces. These mapping and transformation rules are described using Distributed First Order Logic, and KIF is the underlying language used for the PSL Ontology at run time.

In the context of the FUSION Project [2][12], concepts from the data and functional facets are used to annotate web services at development time. Alazeib et al. [2] propose two approaches for semantically-assisted design of collaborative business processes. The manual approach uses previously annotated business process models that define control and data flow; the semi-automatic approach deals with business processes whose control and data flow are defined by planning techniques using behavioral descriptions of the web services and composition goals. In [12], the manual approach is applied for demonstrating the usage of ENIO in facilitating the smooth integration of business processes in a real-world B2B scenario. ENIO is also used at run time, being implemented in OWL-DL. "Up-cast" (from service inputs/outputs to ontology concepts) and "down-cast" (from ontology concepts to service inputs/outputs) transformations are used at run time to perform dynamic data mediation.

Masuya et al. [14] examined, at development time, the schemas of the databases to be integrated and, then classified their contents and semantics under YAMATO-GXO Lite, in a job that they consider to be comparable to database annotation. YAMATO-GXO Lite is implemented in RDF and OWL, and the database annotations are made in these languages to support queries at run time.

In [8], Scherp et al. present several object diagrams to exemplify how the structured knowledge in a concrete scenario of a flooding event can be modeled at development time. The ontology design patterns of the core ontologies are combined and used to model this scenario. The core ontologies are implemented in OWL to enable application communication at run time, but this use is not discussed in the paper.

Paulheim and Probst [15] advocate that application developers should describe their applications in application ontologies, which have to be based on domain ontologies. This is done at development time. In the example presented in [15], they annotate class diagrams with concepts and relations from a domain ontology. Moreover, the ontologies are implemented in F-Logic and the proposed framework uses run-time reasoning to determine possible cross application interactions.

Finally, Treiblmayr et al. [16] present the use of ontologies only at development time. Application concepts are grounded in concepts from DOLCE and from the semantic reference system at development time.

RQ3. Do the studies follow a systematic approach for performing the integration project? (Do they adopt or propose a method or a process model defining activities, inputs, outputs, guidelines, etc.?)

Most of the studied initiatives is developed in an ad-hoc manner [13][14][15][8][16], i.e. they do not follow a systematic approach for integrating applications. The exceptions are the works done in the context of the FUSION Project [2][12], mainly [2].

In [13], Martín-Recuerda only revisits the solution for application integration proposed by Cheng and Law [18] to test the suitability of CSpaces to deal with integration scenarios. Treiblmayr et al. [16] do not follow or define an integration method; concerning methodological aspects, they discuss aspects related to conceptual mapping (using a foundational and a core ontology), and how to semantically describe services. Masuya et al. [14] describe how they conducted the integration initiative, discussing the performed steps, without the pretension of proposing that these steps could be generalized to give rise to a method. The focus of [8] is on designing core ontologies. Concerning the integration initiative, this work only exemplifies how to model the knowledge exchanged by the systems. Finally, Paulheim and Probst [15] propose a framework for application integration on the user interface level by encapsulating the applications in plugins, and pinpoint three steps that are necessary for integrating an application into their framework. In the context of the FUSION Project, Alazeib et al. [2] present two approaches (one manual, another semi-automatic) for semantically-assisted design of collaborative business processes in EAI scenarios. Both can be considered as systematic approaches. Table 1 summarizes this section.

Table 1. Summary of the synthesized data, organized by research question

Study	Approach Employed (RQ1)	Use of Ontology (development/run time) (RQ2)	Integration Approach (RQ3)	Foundational Ontology Adopted
[2]	Indirect Approach	Development Time and Run Time	Systematic	DOLCE & SUMO
[8]	Indirect Approach	Development Time and Run Time	Ad-hoc	DOLCE
[12]	Indirect Approach	Development Time and Run Time	Systematic	DOLCE & SUMO
[13]	Direct Approach	Development Time and Run Time	Ad-hoc	PSL Ontology
[14]	Indirect Approach	Development Time and Run Time	Ad-hoc	YAMATO
[15]	Indirect Approach	Development Time and Run Time	Ad-hoc	DOLCE
[16]	Hybrid Approach	Development Time	Ad-hoc	DOLCE

6 Discussion

In this section, we discuss some relevant points that have arisen from data synthesis and discuss limitations of this review.

The Importance of Foundational Ontologies in EAI. Taking into account the small number of selected studies, we can conclude that the use of foundational ontologies in EAI initiatives is relatively underexplored. Nevertheless, foundational ontologies

have gained attention in the latest years, as corroborated by the publishing year of the studies: 2006 ([13]), 2007 ([2][12]), 2010 ([15]), and 2011 ([14][8][16]).

Foundational ontologies are being used in EAI initiatives, especially by offering fundamental ontological distinctions for solving semantic conflicts. We have observed that most of the selected studies use lightweight versions of domain ontologies, and that these versions are grounded directly or indirectly in foundational ontologies. This indicates that these two kinds of ontologies may be used together in EAI.

Regarding the approaches adopted (direct, indirect, and hybrid approaches), we think that the hybrid approach seems to be the most promising, since it puts together the characteristics of the other two approaches, being, therefore, more embracing. In situations where we do not have domain or core ontologies available to address all semantic heterogeneities in an EAI initiative (indirect approach), applying the ontological distinctions directly to analyze the applications' models may be useful. In sum, we claim that common understanding can be achieved by using domain, task or core ontologies, but also by directly applying foundational ontology. Also, we highlight DOLCE as the predominant ontology, being used by different research groups.

Addressing EAI Initiatives at Different Abstraction Levels. All studies included in the review use ontologies at development time, and most of them also use ontologies at run time (except [16]). At development time, ontologies have being used, among others, for negotiating meaning, for defining mappings and transformation rules, for making annotations, as well as for grounding the concepts adopted by the applications being integrated. At run time, most of the studies implement lightweight versions of the ontologies in a variety of machine-readable languages, such as KIF, RDF, OWL and F-Logics. These versions are applied for query processing, concept matching, and message exchange, among other uses.

The use of ontologies at development time indicates a concern related to a conceptual abstraction level in EAI projects, when meaning is negotiated and a common understanding is pursued. The conceptual level has different concerns if compared to a technological one. However, a clear distinction between these two abstraction levels is not present in the analyzed studies. Analogously to other related research areas, such as Software Engineering, a clear separation of concerns related to conceptual and technological abstraction levels could be further explored in EAI.

Moreover, none of the studies have adopted representation languages that incorporate fundamental ontological distinctions in their constructs. In fact, these studies use ontological distinctions given by the foundational ontology, but these distinctions are not part of the language used for building their models. In our view, aligned to what is advocated in [5], the use of well-founded languages is an important issue for adequately addressing the conceptual abstraction level, at the EAI development time.

The Use of Systematic Approaches in EAI Initiatives. Most of the analyzed studies follow an ad-hoc approach to integration (except [2][12]). In our view, there is a gap regarding methodological issues, mainly if we consider the use of foundational ontologies as part of the integration approach. The gaps we have identified suggest that we need to advance towards the establishment of more prescriptive methods and processes. These should clarify the various integration activities, associating them to goals, inputs, outputs and various quality requirements. Systematic approaches can act as a glue to put together, in an organized way, the aspects raised in the previous

discussion points, in particular structuring the integration process into different levels of abstractions and defining guidelines on how to perform the various integration activities under the guidance of ontological distinctions. We believe this is essential for establishing an engineering approach for ontology-based EAI.

Limitations of This Review. Due to the fact that the study selection and data extraction stages were performed by just one of the authors, some subjectivity may have been inserted. To reduce this subjectivity, a second author was responsible for defining a random sample (including about 35% of the studies) and performing the same stages. The results of each reviewer were then compared in order to detect possible bias. Moreover, terminological problems in the search strings may have led to missing some primary studies. In order to minimize these problems, we performed previous simulations in the selected databases. We decide not to search any specific conference proceedings, journals, or the grey literature (technical reports and works in progress). Thus, we have just worked with studies indexed by the selected electronic databases. The exclusion of these other sources makes the review more repeatable, but possibly some valuable studies may have been left out of our analysis.

7 Final Considerations

This work presented a systematic literature review that aims to analyze the adoption of foundational ontologies in EAI initiatives for facing semantic integration problems and, from that, to identify new paths for future research. After searching and selection, seven studies remained, which were analyzed in order to address three research questions. During data discussion and synthesis, some relevant issues were detected and discussed: (i) despite still being underexplored, in the latest years foundational ontologies have gained attention in EAI initiatives; (ii) the need for conducting EAI initiatives in different levels of abstraction (conceptual and technological) as well as the need to adopt more appropriated representation languages, which consider distinct abstraction levels and ontological grounding; and (iii) the need for systematic approaches in EAI to put together all the aspects aforementioned.

Acknowledgments. This research is funded by the Brazilian Research Funding Agencies CAPES, FAPES/CNPq (PRONEX Grant 52272362/11) and CNPq (Grants 483383/2010-4 and 310634/2011-3).

References

1. Lee, J., Siau, K., Hong, S.: Enterprise Integration with ERP and EAI. Communications of the ACM 46(2), 54–56 (2003)
2. Alazeib, A., Balogh, A., Bauer, M., Bouras, A., Friesen, A., Gouvas, P., Mentzas, G., Pace, A.: Towards semantically-assisted design of collaborative business processes in EAI scenarios. In: 5th IEEE International Conference on Industrial Informatics, Vienna, Austria, pp. 779–784 (2007)
3. Izza, S.: Integration of industrial information systems: from syntactic to semantic integration approaches. Enterprise Information Systems 3, 1–57 (2009)

4. Guarino, N.: Formal Ontology and Information Systems. Formal Ontology and Information Systems 3 (1998)
5. Guizzardi, G.: On Ontology, ontologies, Conceptualizations, Modeling Languages, and (Meta)Models. In: Vasilecas, O., Edler, J., Caplinskas, A. (eds.) Frontiers in Artificial Intelligence and Applications, Databases and Information Systems IV, pp. 18–39. IOS Press, Amsterdã (2007)
6. Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: Sweetening ontologies with DOLCE. In: Gómez-Pérez, A., Benjamins, V.R. (eds.) EKAW 2002. LNCS (LNAI), vol. 2473, pp. 166–181. Springer, Heidelberg (2002)
7. Kitchenham, B., Charters, S.: Guidelines for performing Systematic Literature Reviews in Software Engineering (Version 2.3) - EBSE Technical Report. EBSE-2007-01 (2007)
8. Scherp, A., Saathoff, C., Franz, T., Staab, S.: Designing core ontologies. Applied Ontology 6, 177–221 (2011)
9. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A.: WonderWeb DeliverableD18. Ontology Library, Padova, Italy (2003)
10. SUMO (Suggested Upper Merged Ontology), http://www.ontologyportal.org/
11. Mizoguchi, R.: Yet Another Top-level Ontology: YATO. In: Proceedings of the Second Interdisciplinary Ontology Meeting, Tokyo, Japan, pp. 91–101 (2009)
12. Bouras, A., Gouvas, P., Kourtesis, D., Mentzas, G.: Semantic integration of business applications across collaborative value networks. International Federation for Information Processing (IFIP) 243, 539–546 (2007)
13. Martín-Recuerda, F.: Application Integration Using Conceptual Spaces (CSpaces). In: Mizoguchi, R., Shi, Z.-Z., Giunchiglia, F. (eds.) ASWC 2006. LNCS, vol. 4185, pp. 234–248. Springer, Heidelberg (2006)
14. Masuya, H., Makita, Y., Kobayashi, N., Al, E.: The RIKEN integrated database of mammals. Nucleic Acids Research 870, D861–D870 (2011)
15. Paulheim, H., Probst, F.: Application integration on the user interface level: An ontology-based approach. Data & Knowledge Engineering 69, 1103–1116 (2010)
16. Treiblmayr, M., Scheider, S., Krüger, A., Von der Linden, M.: Integrating GI with non-GI services – showcasing interoperability in a heterogeneous service-oriented architecture. GeoInformatica 16, 207–220 (2011)
17. Grüninger, M.: Ontology of the Process Specification Language. In: Staab, S., Studer, R. (eds.) Handbook on Ontologies, pp. 575–592. Springer, Berlin (2004)
18. Cheng, J., Law, K.H.: Using Process Specification Language for Project Information Exchange. In: Proceedings of the 3rd International Conference on Concurrent Engineering in Construction, Berkeley, USA, pp. 63–74 (2002)

Developing Innovative Services: The Case of the Airport Environment

Sofia Gkika, Katerina Fraidaki, and Katerina Pramatari

Department of Management Science and Technology,
Athens University of Economics and Business,
47a Evelpidon and 33 Lefkados str, Athens, 113 62, Greece
{gkikas,fraidaki,k.pramatari}@aueb.gr

Abstract. Innovative services consist the main point of interest so for academics as for the business world. The basic objective of this paper is to demonstrate the way that innovative services may change the "consumer experience" in the airport environment. In the following lines we describe the procedure of designing innovative services, so for consumers' and visitors' needs as for the needs of the firms which are operated in the airport environment. In order to achieve this goal, we used so the qualitative approach (such as focus groups), as the quantitative approach (collection of 1106 questionnaires).

Keywords: electronic services, innovation, business models.

1 Introduction

Roger's (1998) defines *innovation* as "… the process of introducing new ideas to the firm which result in increased firm performance". Schumpter (1983) confirms the existence of five types of innovation: (1) introduction of new product or the change of the existed one, (2) process innovation, (3) opening of a new market, (4) development of new sources of supply of raw materials and (5) changes in the industrial organization.

Innovation has become the key marketers' tool, in order to ameliorate customers' experience. Being focused on the airport environment, customers are divided into four categories: visitors, travellers, firms and firms' employees. Our basic objective is to design innovative services which will direct in these different groups, taking into consideration their special needs.

The following paper is focused on the first type of innovation; to be more specific we are going to "change the way that existing services operate". The purpose of our study is to be emerged sservices for the case study of a Greek airport. The target groups of these services are on the one hand the visitors and the consumers of the airport and on the other hand the firms and their employees that are activated in the specific environment. We aim to design services which will be bought by the firms in order to be given to their customers while they are situated in the airport. The services will be delivered to the final customer via a platform named CAP- Common Airport Platform. This platform will be managed by the managers of the airport; so the firms as the customers will have access using a specific code. The services that we introduce are: Discount Coupons, Traffic Map, Last Minute Tickets and Time2Gate.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 250–263, 2013.

The following paper is divided in four sections. Section One describes the existent services in the specific context and analyses the methodology of our survey. Section Two describes the methodology that was followed in our research, the related work, the focus groups that revealed and described how the survey was conducted. Section Three analyses the services which will be developed, as well as the proposed business model of each service. Finally, section Four contains a brief discussion.

2 Methodology

In order to design the innovative services, we started our research with exploration of our "customers" (visitors, travelers, firms and firms' employees) needs. In order to achieve this goal, we used to the method of focus groups, as the one of the survey. Finally, we designed the services, describing their full business model.

2.1 Related Work

The first step of our research was the analysis of the context in which we wanted to design the innovative services. Studying the permanent literature we tried to identify the services which have been offered until now, by the airports. Table 1 mentions the basic services. As we may notice they are separated in two categories: services towards firms (B2B Services) and services towards passengers (B2C Services).

Table 1. Existing services

B2B Services	Voice Communication Services
	TETRA Digital Radio Ground to Ground Communication
	IT&T Support Services
	Flight Information Display Service
	Data Communication & Internet Services
	Cabling Services
	Data Centre Services & IT Operations
	Added Value Consulting Services to other Airports and Transportation Operators
	Wireless Access
	Common Use Terminal Equipment (CUTE)
	Common Use Self Service kiosks (CUSS)
Passenger services	Business Center Services
	Internet kiosks
	Card phones
	Wi-Fi (wireless) Internet access

The basic characteristic of this context is the limited time that a passenger/visitor consumes while (s)he visits an airport. Normally, either heavy passengers (people who travel once every three months of more often) or light passengers consume one and half hours in the airport, before they travel; on the other hand airport offers sixty (60) minutes of free Internet in order to fulfil his/her needs.

2.2 Focus Groups

Bruseberg (2001) confirms that the most useful way to understand peoples' needs is *focus groups.* Having studied the existed services, the next steps are on the one hand the investigation of the needs that our target group wants to fulfil the most; and on the other hand the investigation of services that could be considered as innovative. For this reason we conducted five focus groups: two of them consist of managers and employees of the firms and the other three of consumers. Every focus group was consisted of 6 people. In order to choose the members of the groups we used Matthing's questionnaire (2006) (Appendix). Matthing proved that being based on how innovative a person can be considered, we may recognise five types of people (starting from the most innovative to the less one): leaders/ explorers, pioneers, sceptics, paranoids and laggards. The questionnaire was distributed in 59 people of the airport community. We chose 30 people and in every group there was at least one leader/ explorer and no paranoids and laggards.

Having analysed the data of the qualitative approach we identified 40 innovative services (33 services towards passengers/ visitors, 5 services towards firms and 2 services towards the employees of the airport). The services could be categorised into three categories: informational, transactional and hedonic/social services. Their basic characteristic is that they refer either to the advanced version of existing services or to the design of completely new services.

2.3 Survey

In order to evaluate the previous services, we used the quantitative approach, conducting a survey in the airport environment. Because of the limited space of the questionnaire we chose three services from every category. The basic objective of this survey was to understand the attitude towards the services. We measured this construct using a two item scale (very interesting/ not interesting at all and very useful/ not useful at all). At the same time we tried to investigate the basic characteristics of the passengers which will help us to design the Business Models of every service.

We collected 1106 answers. The descriptive characteristics of the sample are: 60% of the people who answered the questionnaire are between 18-35 years old, 51% were female and 30% travel once every six months. Furthermore 50% uses smart phone during his/her visit in the airport and 57% of them uses his/her smart phone in order to obtain new services.

Taking into consideration the above analysis as well as the available technology we will develop the following services: discount coupons, traffic map, last minute tickets, time2gate. In the following sections we will fully describe the services and their proposed business model.

3 Services

The services will be offered to "consumers", firms and employees via the platform CAP; its main objective is to collect all the services and provide to users a more "user friendly" environment.

3.1 Discount Coupons

In the context of the service "Discount Coupons" (Table 2) firms operating inside the airport and at the airport's park, import into CAP's platform offers in order to inform

Table 2. Discount Coupons

Components	**Description**
Value Proposition	Automated provision of offers (for purchases at businesses inside and outside the space of Airport) through electronic means.
Customer Segments	• Registered and non- registered passengers • Registered and non- registered visitors • Airport's staff • Business at airport's area
Channels	The electronic proof of the offer will be provided and / or printed by users through the platform of the Airport, which is accessible via (digital channels): • Smart phones & laptops users • desktops and laptops in kiosks in the Airport • The website of the CAP • site of Airport • The wireless portal Airport • Displays (FIDS, IPTV and CAP dedicated) The offers that are not required to be printed / electronic evidence, the bids will be communicated via display screens of offers via sms to mobile users or e-mail (digital channel). Eventually, users will enjoy the offer made in the shops of cooperating firms in the wider area of the airport (physical channel, indirect collaborative). The electronic proof of the offer will be provided and / or printed by users through the platform of the airport, which is accessible via (digital channels): • Smart phones & laptops users • Desktops and laptops in kiosks in the Airport • The website of the CAP • Site of Airport • The wireless portal Airport • Displays (FIDS, IPTV and CAP dedicated) For offers that do not require printed / electronic evidence, the bids will be communicated via display screens of offers via sms to mobile users or e-mail (digital channel). Eventually, users will enjoy the offer made in the shops of cooperating firms in the wider area of the airport (physical channel, indirect collaborative).
Customer Relationships	Users: passengers, visitors, employees • Service Provider-user • Self-service through automated operations (automation) • Personal service (personal assistance) in the redemption of the offer in partner firms Users: partner enterprises • Service provider-user • Co-creation of value between firms and the airport (value co-creation)
Key Resources	Physical resources: e.g. computers, network infrastructure, printers Intellectual resources: e.g. knowledge about the operation of the service, CMS, algorithms for recommendation systems, etc. Human resources: e.g. staff

Table 2. (*continued*)

	Financial resources: revenues from 'Discount Coupons'
Key Activities	• Planning offers (business plan of the partner firms) • Contracts with affiliated firms • Production: design, implementation and launch of the service, targeted supply available (better recommendations), import supply • Maintenance Services • Production and maintenance of platform disposal services • Production and maintenance of network infrastructure • Approval of an offer • Service planning
Competitive Advantage	• Personalized offers (leadership & innovation) • Airport's brand name • Provision of electronic coupon • Competitive cost (value for money) • Charge for dynamic updates list • Providing billing packages
Key Partners	Airport - Business
Revenue Streams	Charge for business • Free of charge • Charge based on mean's projection o Depending on the means of communication of offers (screens or platform services) o As part of promotional offers (displays, banner, smart phones, Airport's website and CAP, wireless portal Airport) o Depending on bids (or non-peak hours) o For admission offer (personalized or not) to the list o Depending on the target group who chose to promote the offer o Offer sent to the list of e-mails to users, informing users through sms on new offers o Renewal offers o Charge per click o Discount rate (incentives) of the total amount payable o Charge and the service that is shown (eg, Way-Finding, Social Networking) o Construction costs for promotional material • Fee based on offer's sales o Based on the number of sales due to the offer (per purchase) o A rate of sales Charge for passengers, visitors, staff • Charge per use of offer • One-time charge for the use of offers Free distribution
Cost Structure	• Purchase of hardware and software • Maintenance of hardware and software • Staff's salaries • Cost of ownership messages • Fixed monthly operating costs

the public that have access to the platform through available technological resources (Wi-Fi, SMS, 3G, Airport Campus LAN, etc.). The offers will be provided on a list. Recipients of this service will be all employees, passengers, visitors of the airport as well as residents of neighbouring municipalities. The service will be available through CAP's available communication channels and different types of devices such as laptop, smart phone or tablet PC within the airport or users with Internet access (to region around the airport). Moreover, "Discount Coupons" will be available even

to people who do not use or do not have any technological means, through the promotion of the offers on monitors located in areas of increased traffic like staff's restaurants or Metropolitan EXPO during an exposure.

The offers that will be provided through CAP will be inserted from the firms. The system can be accessed quickly and easily by any employee of the firm, if (s)he knows the offer's characteristics, the means the offers will project, the time of display, etc. This process is automated so that no special skills or computer knowledge is required.

The actors that act in the service "Discount Coupons" will be the authorized firm's employee that will insert the offer to the Platform; will select the means that the offer will be provided, will check the cost and send the offer to the airport's employee for approval. Subsequently, the airport's authorized employee will be informed about the offer and its cost then will decide whether (s)he will approve or reject it. If (s)he will approve it, offer will be activated, by enabling the provided means. Otherwise, (s)he will inform the airline for the reasons of rejection. The users will be informed for the offer through the means that the firm selected to provide them. In the case of coupons, they download them to their personal devise the coupon's barcode and show it to the cashier while they are paying. The cashier scans the barcode and sends it to the list with the used coupons.

The most important element of the offers' display is personalization. This will be done according to the profile's data that a user entered while registered and / or according to his/her personal history based on the earliest use of coupons. This approach involves registered users and staff of the airport. For non-registered users, the offer proposals will provide personalized via the service "Time2Gate". In this service, the personalized offers depend on the "route" the user seems to follow to the gate of his flight, providing the offers from shops that seem to pass in order to benefit from them and other criteria.

The characteristics that combine the competitive advantage of this service is not only the personalized offers according to user's needs but also the brand of the airport, which gives greater glory to the service. Also, dynamic pricing for the renewal of the list with the offers and coupons, depends on list's amount of visitors when the firm wishes to make the renewal. One more characteristic is the provision of special cost packages for an offer that meet business needs. Moreover, even if there are other services providing this kind of service the advantage of this one is the fact that the user when works or visits the airport has a particular mood which is connected with the airport's environment, so if (s)he knows about this service (s)he will prefer this one.

3.2 Traffic Map

This service "Traffic Map" (Table 3) will also be available through the CAP, and it will give users the opportunity to learn about the traffic to and from the airport. Focus groups revealed that it is a very important service, which is very high at passengers' preferences, since information about traffic offers great convenience to anyone that is interested on accessing the airport grounds. This service refers to everyone who wants to go to the airport or leave from the airport. The user can be informed for the alternative access routes to the airport, the total time required for each route and the traffic along the route. The data of the route provided by the National Police

Table 3. Traffic Map

Components	**Description**
Value Proposition	Estimation of the transit time to/from the airport via the default route and information on whether there are any incidents along the way. Moreover, users may choose to be provided personalized itineraries according to their needs and only registered users have the option to save the route.
Customer Segments	• Registered and non- registered passengers • Registered and non- registered visitors • Airlines
Channels	The service will be accessible via (digital channels): • Smart phones, laptops and users of Tablet PC, • CAP's website • Airport's website and • IPTV monitors
Customer Relationships	Users: passengers, visitors, employees • provider-user • self-service through automated operations (automation)
Key Resources	Physical resources: e.g. computers, network infrastructure Intellectual resources: e.g. the brand of the AIRPORT, the required knowledge, data, algorithms assessment years, the software that implements the service Human resources: e.g. staff Financial resources: the use of service
Key Activities	• Production: design, implementation and launch of the service • Maintenance Services • Ability to personalize, save routes, notification • Dynamic Update • Production and maintenance of platform's disposal services • Production and maintenance of network infrastructure • Calculation of transit time • Update on events affecting the route.
Competitive Advantage	• Airport's brand name • Dynamic information about the events that affects the way • Ability for customization and / or storage paths • Ability for notification and opportunity for dynamic update of users
Key Partners	Airport– Provider (real time data)
Revenue Streams	Charge for passengers, visitors • Free distribution • One-time charge for the use of the service • Charge per text message to mobile Charge for airlines • Free distribution • Single charging for the use of the service Charge for businesses • View offers by service 'Discount Coupons'
Cost Structure	• Hardware and Software equipment • Maintenance of Hardware and Software • Human resources' salaries • Cost of acquiring sms packages

(Local Traffic Department), according to the press issued by traffic and other sources that can provide that kind of information.

The user selects the starting point, the means of transportation to the airport and the estimated time is calculated automatically. Then, the path shown on the chart and is

updated for any traffic incidents. If the user is registered, is given the option to save the above route.

The user is able to estimate the required time to go to the airport through default routes. Additionally, there is the possibility of providing personalized itineraries depending on user needs. For example, routes are customized to user needs in order to select either the fastest route or a route with sightseeing. Finally, only the registered users can save the route they chose or receive sms if there is a change in the time that they had originally assumed, because of a potential accident.

The characteristics that build the competitive advantage of the service "Traffic Map" is that the estimated time that displayed to users is based on real information about any potential problems on the route to the airport. Also, it may personalize and / or store routes that a user chose and gives the opportunity for notice and dynamic update of users.

3.3 Last Minute Tickets

The aim of this service is to provide cheap air tickets for flights departing the next one to four hours (Table 4). The service will be available to everyone who uses the CAP which is separated in two different levels depending on whether the user is registered or not. In the process of registering a new user can choose if (s)he wants to receive information for promotions on the service either via SMS or via e-mail. Registered users will have the privilege to be informed through SMS or e-mail, one or two hours before the non-registered users in order to keep them on a priority over other users. Because of the fact that the registered users have the privilege of being informed before non-register users, they may be located outside the airport before receiving the information and go to the airport only if (s)he is interested in it. It should be mentioned that the tickets will be "aller-retour".

The actors of the service "Last Minute Tickets" are the authorized airline's employee that insert the offer to the platform, select the means that the offer will be provided, check the cost and send the offer to the airport's employee for approval. Subsequently, the airport's authorized employee informed about the offer and the cost then decides whether to approve or reject it. If (s)he approves it, the offer is activated through the enable of the means that was decided to be projected. Otherwise, (s)he informs the airline for the reasons of rejection. Finally, the users, if they are registered, they are informed earlier than the non-registered user through an sms or an e-mail, or from the airports displays and moves to the airline's desks in order to purchase the ticket.

The most important element of this service is the presence of innovation, as there is no similar service, at least in national data, providing cheaper tickets to users. The user's access to information and booking tickets will be easy and the process will be completed quickly.

The tickets that will be provided through CAP will be inserted from the airlines. This can be done quickly and easily from any employee of the firm, if (s)he knows the login information to the platform, the number of tickets that will be provided, the projection part of tickets, time of display, etc. This process is automated so that not requires any special skills and computer knowledge / technology to use.

Table 4. Last Minute Tickets

Components	**Description**
Value Proposition	Automated allocation of financial tickets (aller-retour), for flights departing in 1-4 hours from the time the offer will be displayed on the electronic means.
Customer Segments	• Registered and non- registered passengers • Registered and non- registered visitors • Airlines partners • Airport's staff
Channels	Notification for registered users through sms (digital channel). Users are informed through the platform, which is accessible via (digital channels): • Smart phones, Laptops & Tablet PC • Desktops, Laptops & Tablet PC's at kiosks within the Airport • FIDS screens and IPTV in Airport. Automatic import and tickets' renewal from airlines through the platform of the Airport (digital channel). Eventually, people will buy tickets directly from the airlines' desk (physical channel, indirect collaborative).
Customer Relationships	Users: airlines • Service provider-user • Co-create value between airlines and Airport (value co-creation) Users: passengers, visitors, employees • Service provider-user • Self (self-service) through automated operations (automation). • Personal service when buying tickets on airlines (personal assistance)
Key Resources	Physical resources: e.g. computers, network infrastructure equipment for the disposal of the service through transmission service Intellectual resources: e.g. knowledge about the operation of the service, databases, software, algorithms Human resources: e.g. staff Financial resources: revenues from 'Discount Coupons'
Key Activities	• Design for provided offers (business plan of the partner airlines) • Contracts with the partner airlines • Production: design, implementation, service delivery. • Maintenance service. • Production and maintenance of platform disposal services. • Production and maintenance of network infrastructure.
Competitive Advantage	• Personal electronic notice to registered users (leadership & innovation) • Information through digital channels for all users • Airport's brand name • Competitive cost (value for money) • Dynamic renewal of fee list • Provision of preferential packages
Key Partners	Airport – Airlines
Revenue Streams	Charge for business • Free of charge • Charge based on mean's projection o Depending on the means of communication of offers (screens or platform services) o As part of promotional offers (displays, banner, smart phones, Airport's website and CAP, wireless portal Airport) o Depending on bids (or non-peak hours) o For admission offer (personalized or not) to the list o Depending on the target group who chose to promote the offer

Table 4. (*continued*)

	○ Offer sent to the list of e-mails to users, informing users through sms on new offers ○ Renewal offers ○ Charge per click ○ Discount rate (incentives) of the total amount payable ○ Charge and the service that is shown (eg, Way-Finding, Social Networking) ○ Construction costs for promotional material • Charge based on offer's sales ○ The number of sales due to the offer (per purchase) ○ A rate of sales Charge for passengers, visitors, staff • Charge per use of offer • One-time charge for the use of offers • Free distribution
Cost Structure	• Purchase of Hardware and Software • Maintenance of Hardware and Software • Staff's salaries • Cost of ownership messages • Fixed monthly operating costs

The components that create competitive advantages that are met in this service is that the service is innovative and all, potentially, customers will wish to know and to use it even once. All processes are automated and no effort is required for the use of this service. Last but not least, the competitive costs (value for money), makes it even more attractive service to customers.

3.4 Time2Gate

This service will have as main purpose to provide dynamically information to user (passenger) for the estimated time to go to the terminal gate and is addressed to all CAP's users (Table 5). It should be mentioned, that this is the only service that is applied under the condition that the user is in the airport area. This service is mainly addressed to passengers but the airlines are also considered as users. Additionally, there is the ability to be provided dynamically information on travel time for the requested departure gate through some standard points that are declared as Internet access points (kiosks) and Information Offices (Information Desks) located in the Main Terminal Building.

The passengers may have access to this service through their smart phone or via computer or a kiosk. Another option is the access via the information desk of an employee. If the passenger uses smart phone, the position will be identified (if (s)he permits it), with the insert of the flight's number in order to be defined the departure gate of the flight and then will display the estimated time at each point in the process. Finally, the map of the route to the departure gate is displayed. In case where a passenger will go to a kiosk, the location of the kiosk will be identified automatically and the estimated time will be calculated and then the user will provide the flight's information. Finally, if (s)he will go to the information desk, the information will be provided through the employee by informed him/her about the code of flight and (s)he then informs the passenger about the estimated time. The employee of the airport will

Table 5. Time2Gate

Components	**Description**
Value Proposition	Dynamic, real-time (real-time) provision of information to passenger for the required time to go to the boarding gate of his/her flight.
Customer Segments	• Registered and non- registered passengers • Registered and non- registered visitors
Channels	The users may use this service through: • Smart phones & laptops users • Desktops and laptops in kiosks in the Airport In case of a substantial change in the next years points the user will be informed with sms for new hours (digital channel)
Customer Relationships	Users: businesses • Service provider -user • Co-create value between airlines and the Airport (value co-creation) Users: airlines, passengers, visitors • Self - service through automated processes (automation). • Service provider –user
Key Resources	physical resources: e.g. computers, network infrastructure, sensors for detecting user location and transit time calculation intellectual resources: e.g. knowledge about the operation of the service, algorithms / methods of positioning algorithms for calculating the access time at the gate Human resources: e.g. staff Financial resources: revenues from 'Discount Coupons'
Key Activities	• Selection and implementation of algorithms • Production: design, implementation and launch of the service • Maintenance Services • Production & maintenance platforms services • Production and maintenance of network infrastructure
Competitive Advantage	• Calculation of the gate transit time (combination of Automatic Identification and assessment of service time in queues and checkpoints) • Dynamic Update • Innovation, Leadership • Airport's brand name
Key Partners	• Business • Airlines
Revenue Streams	Charge for business • Charge for viewing offers through service 'Discount Coupons' • Charge to promote businesses • Economy charge packets Charge for users • One-time charge for sending sms for information in case of a substantial change in the years following points • One-time charge for using the service to automatically detect the user's position • One-time charge of mere user of the service (without automatic user identification) • Free of charge Charge for airlines • One-time charge for the rights to use the service • Depending on the times you used the service
Cost Structure	• Purchase of hardware and software • Maintenance of hardware and software • Human Resources Salaries • Equipment for positioning systems and calculation queue at checkpoints • Cost of acquisition sms packages

log in to the application, and then (s)he will insert the code of the flight and where they are located. Then the time will be estimated and the passenger will be informed about the time that is required in order to go to the gate.

Such a service is used for the first time in the airport area. Furthermore, the user access and use of service will be easy, without requiring specialized knowledge. Finally, the service will be followed by a strong brand name that inspires confidence and trust to people.

The calculation of the gate transition time (a combination of Automatic Identification and assessment of service time in queues and checkpoints) is an innovative service. The dynamic information of the user in case there will be a problem or the sudden existence of a long queue that will increase the estimated time to gate, are innovative aspects and create competitive advantage.

4 Discussion

The purpose of our study was to create a platform with services for the passengers, the visitors the firms and their employees of an airport. Through a survey that was conducted with focus groups and questionnaires, four services identified as the most desirable. These services as we described above in detail, are Discount Coupons, Traffic Map, Last Minute Tickets and Time2Gate.

Our research has two important managerial implications. The former is the way that the airport may help the firms to fulfil their customers' needs with a better way, recognising their personal characteristics. Recommender systems become one of the basic parts of the delivery of our services. Being based on this managerial implication, the next objective of our research is the investigation of the way that we may design a more intelligent system; a system which will be able to evaluate the data and propose directly solutions so to the firms as to their customers. The latter, the (technical) way that the recommender systems should be designed in order to achieve higher accuracy to their results.

Acknowledgments. Project Action "COOPERATION 2009" entitled "Common Airport Platform- CAP" code 09SYN-72-799 and sponsored by GSRT. / EYDE-RTDI - Ministry of Education, Lifelong Learning and Religious Affairs'.

References

Adamic, L., Adar, E.: How to search a social network. Social Networks 27(3), 187–203 (2005)

Bhattacherjee, A.: Individual trust in online firms: Scale development and initial test. Journal of Management Information Systems 19(1), 211–241 (2002)

Chakravarty, A., Liu, Y., Mazumdar, T.: The Differential Effects of Online Word-of-Mouth and Critics' Reviews on Pre-release Movie Evaluation. Journal of Interactive Marketing 24(3), 185–197 (2010)

Cheung, C.M., Lee, M.K.: Understanding consumer trust in Internet shopping: A multidisciplinary approach. Journal of the American Society for Information Science and Technology 57(4), 479–492 (2006)

Corbitt, B.J., Thanasankit, T., Yi, H.: Trust and e-commerce: a study of consumer perceptions. Electronic Commerce Research and Applications 2(3), 203–215 (2003)

D'Souza, C., et al.: Green decisions: demographics and consumer understanding of environmental labels. International Journal of Consumer Studies 31(4), 371–376 (2007)

Dellarocas, C.: The digitization of word of mouth: Promise and challenges of online feedback mechanisms. Management Science 49(10), 1407–1424 (2003)

Dellarocas, C., Zhang, X., Awad, N.F.: Exploring the value of online product reviews in forecasting sales: The case of motion pictures. Journal of Interactive Marketing 21(4), 23–45 (2007)

Deutsch, M.: Trust and suspicion. Journal of Conflict Resolution 2(4), 265–279 (1958)

Duan, W., Gu, B., Whinston, A.B.: The dynamics of online word-of-mouth and product sales–An empirical investigation of the movie industry. Journal of Retailing 84(2), 233–242 (2008)

Einwiller, S., Geissler, U., Will, M.: Engendering trust in Internet businesses using elements of corporate branding. In: Proceedings of the 2000 Americas Conference on Information Systems (AMCIS 2000), pp. 733–739 (2000)

Friedman, B., Khan, P.H., Howe, D.C.: Trust online. Communications of the ACM 43(12), 34–40 (2000)

Galletta, D.F., et al.: Social influence and end-user training. Communications of the ACM 38(7), 70–79 (1995)

Gauri, D.K., Bhatnagar, A., Rao, R.: Role of word of mouth in online store loyalty. Communications of the ACM 51(3), 89–91 (2008)

Gefen, D., Straub, D.W.: Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. Omega 32(6), 407–424 (2004)

Gill, H., et al.: Antecedents Of Trust: Establishing A Boundary Condition For The Relation Between Propensity To Trust And Intention To Trust. Journal of Business and Psychology 19(3), 287–302 (2005)

Grabner-Kräuter, S., Kaluscha, E.A.: Empirical research in on-line trust: a review and critical assessment. International Journal of Human-Computer Studies 58(6), 783–812 (2003a)

Grabner-Kräuter, S., Kaluscha, E.A.: Empirical research in on-line trust: a review and critical assessment. International Journal of Human-Computer Studies 58(6), 783–812 (2003b)

Hann, I.H., et al.: Online information privacy: Measuring the cost-benefit trade-off. In: 23rd International Conference on Information Systems (2002)

Harrison McKnight, D., Choudhury, V., Kacmar, C.: The impact of initial consumer trust on intentions to transact with a web site: a trust building model. The Journal of Strategic Information Systems 11(3-4), 297–323 (2002)

Hu, N., Zhang, J., Pavlou, P.A.: Overcoming the J-shaped distribution of product reviews. Communications of the ACM 52(10), 144 (2009)

Katz, J.E., Rice, R.E.: Social consequences of Internet use: access, involvement, and interaction. MIT Press (2002)

Kim, D.J., Ferrin, D.L., Rao, H.R.: A study of the effect of consumer trust on consumer expectations and satisfaction. In: Proceedings of the 5th International Conference on Electronic Commerce (ICEC 2003), Pittsburgh, Pennsylvania, pp. 310–315 (2003a), http://portal.acm.org/citation.cfm?id=948046 (accessed December 3, 2010)

Kim, D.J., Ferrin, D.L., Rao, H.R.: A study of the effect of consumer trust on consumer expectations and satisfaction. In: Proceedings of the 5th International Conference on Electronic Commerce (ICEC 2003), Pittsburgh, Pennsylvania, pp. 310–315 (2003b), http://portal.acm.org/citation.cfm?id=948046 (accessed December 4, 2010)

Kim, D.J., Ferrin, D.L., Rao, H.R.: A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. Decision Support Systems 44(2), 544–564 (2008)

Kim, D.J., et al.: A multidimensional trust formation model in B-to-C e-commerce: a conceptual framework and content analyses of academia/practitioner perspectives. Decision Support Systems 40(2), 143–165 (2005)

Koufaris, M., Hampton-Sosa, W.: Customer trust online: examining the role of the experience with the Web-site. Department of Statistics and Computer Information Systems Working Paper Series, Zicklin School of Business. Baruch College, New York (2002)

Koufaris, M., Hampton-Sosa, W.: The development of initial trust in an online firm by new customers. Information & Management 41(3), 377–397 (2004)

Kuan, H.H., Bock, G.W.: Trust transference in brick and click retailers: An investigation of the before-online-visit phase. Information & Management 44(2), 175–187 (2007)

Lee, J., Lee, J.: Understanding the product information inference process in electronic word-of-mouth: An objectivity-subjectivity dichotomy perspective. Information & Management 46(5), 302–311 (2009)

Lee, M.K., Turban, E.: A trust model for consumer Internet shopping. International Journal of Electronic Commerce 6(1), 75–91 (2001)

Levitan, L.C., Visser, P.S.: Social network composition and attitude strength: Exploring the dynamics within newly formed social networks. Journal of Experimental Social Psychology 45(5), 1057–1067 (2009)

Liu, C., et al.: Beyond concern–a privacy-trust-behavioral intention model of electronic commerce. Information & Management 42(2), 289–304 (2005)

McCole, P., Ramsey, E., Williams, J.: Trust considerations on attitudes towards online purchasing: The moderating effect of privacy and security concerns. Journal of Business Research 63(9-10), 1018–1024

Oliver, R.L., Burke, R.R.: Expectation Processes in Satisfaction Formation. Journal of Service Research 1(3), 196–214 (1999); Olson, J.S., Olson, G.M.: i2i trust in e-commerce. Communications of the ACM 43(12), 41–44 (2000)

Ou, C.X., Sia, C.L.: To trust or to distrust, that is the question. Communications of the ACM 52(5), 135 (2009)

Pustejovsky, J.E., Spillane, J.P.: Question-order effects in social network name generators. Social Networks 31(4), 221–229 (2009)

Einwiller, S., Will, M.: The Role of Reputation to Engender Trust in Electronic Markets (2001), http://www.communicationsmgt.org/modules/pub/download.php?id=communicationsmgt-11 (accessed December 4, 2010)

Senecal, S., Nantel, J.: The influence of online product recommendations on consumers' online choices. Journal of Retailing 80(2), 159–169 (2004)

Sheppard, B.H., Sherman, D.M.: The Grammars of Trust: A Model and General Implications. The Academy of Management Review 23(3), 422–437 (1998)

Subramani, M.R., Rajagopalan, B.: Knowledge-sharing and influence in online social networks via viral marketing. Communications of the ACM 46(12), 300 (2003)

Teo, T.S., Liu, J.: Consumer trust in e-commerce in the United States, Singapore and China. Omega 35(1), 22–38 (2007)

Walczuch, R., Lundgren, H.: Psychological antecedents of institution-based consumer trust in e-retailing. Information & Management 42(1), 159–177 (2004)

Creating Knowledge within a C-Business Context: A Customer Knowledge Management View

Apostolos Skotis, Ioannis Katsanakis, Aristomenis Macris, and Michael Sfakianakis

Department of Business Administration, University of Piraeus, Greece
{askotis,ikatsana,arism,msfakian}@unipi.gr

Abstract. Recent advances of virtual networking technologies are gradually forcing companies to focus their knowledge management efforts to external knowledge resources, in order to complement their existing knowledge bases, find expertise, but also harness collective intelligence that is dynamically produced in the virtual environment. Access, exchange and co-creation of customer knowledge is of central importance for companies in this context, as customers who take advantage of Web 2.0 connectivity and social networking tools are gaining importance as competitive and cooperative knowledge actors in companies' C-Business value networks. In this paper the authors attempt to cover important issues concerning customer knowledge flows between companies and customers through virtual interaction and the important factors that determine value-adding relationships of cooperation with customers for effective knowledge co-creation. They emphasize the need for the formation of a strategic co-opetition perspective for managing these relationships. In this direction, the authors present a theoretical framework that describes Customer Knowledge Management within a C-Business context.

Keywords: Knowledge Creation, C-Business, Co-opetition, Customer Knowledge Management.

1 Introduction

Nowadays, the unsteady and fragile economic and business environment strongly imposes new approaches for firms in order to achieve their strategic goals and ensure their long-term survival. Many of them try to develop and establish better relationships with their suppliers or customers, and in general with their stakeholders, in order to combine some of their resources and capabilities to create competitive advantages over their competitors (Najmaei & Sadeghinejad, 2009; Katsanakis & Kossyva, 2012). The development of co-opetitive (cooperative and simultaneously competitive) relationships between firms and stakeholders could help them create competitive advantages (Afuah, 2000; Bengtsson & Kock, 2000; Dagnino & Padula, 2002; Gnyawali & Park, 2009; Ritala & Hurmelinna-Laukkanen, 2009; Wagner et al., 2010). These co-opetitive relationships could be developed with the support of Information and Communication Technology (ICT) infrastructure and the use of virtual networks within a C-Business context (Katsanakis & Kossyva, 2012).

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 264–277, 2013.

C-Business is the integration of E-Business, knowledge management and collaboration technologies. It is a form of collaboration between firms and their stakeholders in which they form an alliance with the use of Information and Communication Technologies (ICT) in order to share and create knowledge (Katsanakis & Kossyva, 2012). C-Business creates dynamic collaborations and transforms firms' information and knowledge base into a computer-based framework to support individualized access to potentially all participants within the alliance (Holsapple & Singh, 2000; Kim et al, 2006). This computer-based framework is primarily based on web-based technologies and services which enable firms to collaborate and share knowledge (Zhang & Deng, 2008).

The purpose of this paper is to examine how external knowledge, and especially customer-generated knowledge, can be used and managed by firms within a C-Business context. Therefore, in this paper the authors attempt to cover important issues concerning customer knowledge flows between companies and customers through virtual interaction and the important factors that determine value-adding relationships of cooperation with customers for effective knowledge co-creation. Moreover, the authors present a theoretical framework on how customer-generated knowledge can be managed within a C-Business context and how customer and firms can collaborate to create knowledge.

2 Knowledge Creation within a C-Business Context

According to Katsanakis & Kossyva (2012), firms within and beyond the value chain, create a virtual network in which, with the support of collaborative technologies, are cooperating in order to create value. This virtual value network includes different kinds of collaborating actors within a certain industry; direct and indirect competitors, suppliers and buyers, and highlights the interaction and interdependence among them. Within this virtual network a common knowledge base is created, where each actor will be able to use and exploit this knowledge for individual purposes. The value is created via knowledge exchange, creation and transfer from the collaboration established within this virtual network. This collaboration is originated from the firms' co-opetition strategy, enhanced with the value creation process (Katsanakis et al, 2011; Kossyva et al, 2011).

As shown in Figure 1, firms and their stakeholders are able to create new knowledge through the sequence of the processes of identification, acquisition, dissemination, transfer, application and creation, which will lead to value creation. The aim of these collaborative relationships is the development of win-win situations for all collaborating actors and the creation of value for customers. The knowledge created is collective, and essentially represents the overall value created from the collaboration of the participating in the virtual network firms. After the creation of the collective knowledge, originated from the collaboration between the participating in the virtual network firms, collaborating actors can utilize the generated value for the creation of individualized value. This depends on each actor's knowledge absorption capacity and knowledge base, which varies from firm to firm (Katsanakis & Kossyva, 2012; Kossyva et al, 2011).

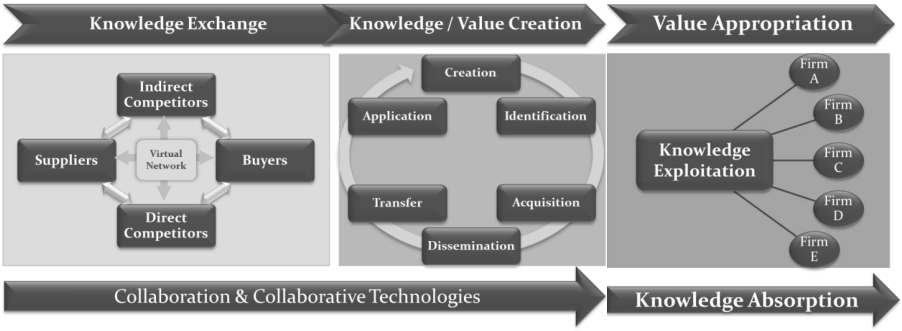


Fig. 1. Knowledge and Value Creation within a C-Business context (Katsanakis & Kossyva, 2012)

3 The Growing Importance of External Knowledge Resources

Externally generated knowledge is considered to be an important source of competitive advantage for a firm, provided that the latter has the potential capacity to acquire, assimilate and exploit it (Zahra & George, 2002). Modern firms which are able to develop systems that increase external knowledge flows, can dramatically improve their ability to innovate, cope with turbulent market conditions and introduce novel ideas, experiencing advantages in terms of speed, cost and quality (Rigby & Zook, 2002). The governing structures a company uses in order to mix, re-combine and find complementarities between internal and external knowledge define processes of knowledge creation and growth and radically affect a company`s knowledge creating competences (Patrucco, 2009).

In this extended knowledge context, organizations face the challenge of widening the range as well as the efficiency of access to distributed external capabilities and knowledge sources. Therefore, established information technology and knowledge infrastructures should provide, among others, solutions for better monitoring and integration of external knowledge resources as well as properly coordinated interaction platforms, promoting integration of complementary internal and external competences, or, more broadly, access to new information, idea exchange, and expertise through electronic networks and in social interaction virtual spaces (Maier, 2005; Consoli & Patrucco, 2010; Wasko & Faraj, 2005).

Following the introduction and rapid growth of the internet and virtual networking technologies, firms have been able to facilitate significantly the productive use of external knowledge and learning resources, as well as knowledge creation and dissemination processes (Eng, 2004), while the more recently advanced Web 2.0 networking and collaboration technologies have paved the way for a shift from traditional knowledge-push to more dynamic, social, people driven learning models. Providing a wide range of virtual networking and collaborative tools, these technologies have greatly enhanced knowledge creation capabilities, driven by the need of connecting different and remote knowledge sources and of harnessing collective intelligence as it is produced in virtual interaction spaces (Chatti et al., 2007).

4 Customer Knowledge

Recent Knowledge Management (KM) literature makes clearly evident that emphasis shifts from the interior, mainly concerning mobilization and access to organizational knowledge, to the exterior of the organization, as the importance of external knowledge sources is increasing. In this manner, a new stage of knowledge management evolution, or, in other words, of KM maturity has been introduced, representing all aspects of external knowledge resources and corresponding knowledge processes and activities (Koenig, 2008). This shift of knowledge locus to outside corporate boundaries demands new views and considerations for accessing customer knowledge, supporting customer integration (Gurgul et al., 2002). Organizations, in their effort to access customer insight, can develop strategies that support extended dialogues with customers, creation of customer knowledge communities and other customer knowledge activities.

Customer knowledge can create value for organizations by contributing to the formation and improvement of planning and operation activities, while it also supports the organization in understanding latent customer needs and expectations. So, access to customer knowledge resources helps a company to form marketing and sales strategies, predict future needs, identify customer decision making crucial factors, and even enhance new product and service development (Lesser et al., 2000).

A significant part of customer knowledge streams emerges from the use of new technologies of social networking or, more generally, virtual interaction technologies that have the ability to deal with various aspects of customer knowledge and experience, even those that are of more tacit nature (Liberona et al., 2012) which are considered to be valuable as a genuine source of innovation and knowledge creation (Polanyi, 1966; Nonaka & Takeuchi, 1995). In this context, customer is now seen as a knowledgeable entity, rather as a passive recipient of products and services. He is an active co-creator of value. Corporate efforts to deal with acquiring, sharing and expanding knowledge residing in customers, to both company and consumer benefit, has brought significant attention to Customer Knowledge Management (CKM). Its main objective is joint value creation by interacting with customers (Gibbert et al., 2002).

5 Customer Knowledge Management

The introduction of Customer Knowledge Management (CKM) came from the need to acquire knowledge created during social interactions that take place between employees and customers, that traditional Customer Relationship Management (CRM) could not identify and acquire, viewing the customer as an agent who seeks but also holds knowledge valuable to the company, thereby establishing a two-way knowledge flow relationship (Murillo & Annabi 2002). CKM is usually portrayed as the integration between CRM and KM, in the sense that KM tools and processes provide the necessary means to support customer knowledge exchange for improving customer service and ongoing relationships (Rollins et al, 2005).

However, a universally accepted definition of CKM does not exist, but it could be generally defined as a continuous process of generation, dissemination and usage of customer knowledge, both inside a company and between a company and its customers (Buchnowska, 2011), but also between customers (C2C) (Zanjani et al., 2008). Gibbert et al. (2002) have identified different styles of CKM, varying by the intensity of interaction, main objectives and knowledge types exchanged. They indicate that knowledge co-production taking place in joint innovation development initiatives as well as communities of creation can facilitate the identification and exchange of tacit knowledge, using strategies such as developing close relationships with lead customers that carry expert knowledge or putting together customer groups that interact with both the company and each other.

A number of different CKM taxonomies have been introduced, but the prevailing distinction is between the following four CK types (Gebert et al., 2003), (Smith and Mc Keen, 2005):

- **Knowledge about Customers:** Considered to be a type of knowledge useful for knowing the customer better, accumulated to identify motivations and by analyzing purchasing activity, connections, history, requirements and expectations (Gebert et al., 2003), future desires and financial capability (Salomann et al., 2005). Acquisition is achieved in a passive way, meaning that it is not the product of active interaction with customers (Buchnowska, 2011).
- **Knowledge from Customers:** Represents knowledge of customers about products, suppliers and market trends. If used through appropriate mechanisms, it can can be a valuable source that can lead to new idea generation, product and service continuous improvements as well as new product and service development (Gebert et al., 2003), (Salomann et al., 2005).
- **Knowledge for Customers:** The type of knowledge regarding customers` specific knowledge needs concerning products, suppliers and markets. It can be acquired not only by company resources, but also from other customers, information consulting institutes and competitors. Also, it has been proposed that it represents the integration of the two aforementioned customer knowledge types (Gebert et al., 2003), (Buchnowska, 2011).
- **Knowledge with Customers (Co-Creation):** Derived from productive dialogue and cooperation of a company with its customers, by using KM procedures and tools to facilitate interaction and knowledge creation (Smith and McKeen, 2005).

Customer knowledge co-creation perspective demands the development of collaboration with customers for joint value creation. It also views customers as members of a wider network, able to extract business value and playing a much more active, multiple role as collaborators, co-developers and competitors (Lawer, 2005).

We use the taxonomy of knowledge about, for, from and with customers to describe the knowledge flows that take place between the company and its customers, either individually or collectively through the use of virtual networks and especially Web 2.0 collaboration technologies. Customers, as knowledge holders or seekers are strengthening their position as players in the virtual environment, taking advantage of

the mass usage of these dominant new technologies by users - consumers. Companies are obligated to view customers` position in the coopetitive value network as more active, energetic and value adding of higher potential. In terms of competitive or cooperation ability, customers are gradually equaling companies` competitors, only with a different set of costs and benefits guiding their motives. Next, we draw several concepts from relevant literature in order to support this position.

6 How New Virtual Network Technologies Affected Customers` Role and Position

6.1 Virtual Network Technologies

The introduction of the Web 2.0 concept came as an attempt to describe a whole new range of new principles and techniques that link sites and users. Web 2.0 technologies have brought a new generation of services that evolved and grew along with users` massive interactions, cooperations and content contributions. The vast amount of new networking and connectivity tools along with the growing collective intelligence created by user generated content (UGC) allow for mass collaboration and online communities formation. This has allowed firms using Web 2.0 tools not only to rapidly learn more about their customers, but also to be able to provide personalized information support and, more importantly, incorporate them as active co-creators of knowledge and co-developers of new products and services (Boselli et al., 2008, Sigala, 2009). Customers now use all available media and communication channels to interact, create and share content, join and participate in social networks and build relationships with other customers (Hennig-Thurau, 2010). Virtual interactions supported by social software promote more efficient ways of bringing customers closer to companies, by facilitating CKM processes and conversions between tacit and explicit customer knowledge (Zhang, 2011). For example, online communities of practice can constantly support knowledge creation activities through technological tools, collective reflection and usage scenarios. Generally, virtual communication tools variety and diversity is crucial to enhancing user tacit knowledge explication, which is the essence of most knowledge creation efforts. (Hemetsberger & Reinhardt, 2004).

6.2 Customers` Enhanced Role and Position

As the internet became the universal information gathering medium, it affected customers by making them more intelligent, in the sense that they could have broader and easier access to information as well as numerous communication channels to share that information (Greenberg 2009). Gradually, several types of “virtual” customers emerged. The “digital consumer”, is described as the consumer who searches for products and product information online. This has urged businesses to provide more in-depth information, comparison and search options and has forced them to consider new ways of understanding consumers` needs and preferences and forming valuable partnerships with consumers. The digital consumer is also an active

creator of online content for brands and products. His activity covers a wide range of created digital content from product reviews and information creators to online designing of customized products (Fiore, 2008).

The growing importance of the customer who acts online is also highlighted by the emergence of the term electronic word-of-mouth (EWOM), all the articulations of customers` comments and reviews about products, services and firms which take place online and circulate through digital information and communication channels. In text or multimedia form, EWOM is rapidly forwarded to other users – customers and has forced firms to understand its determinants and effects on customer decision making. Given that online information access and exchange costs very low, EWOM can develop very large scalability and create new market dynamics. It has also urged companies to develop effective response strategies to manage negative EWOM. (Hennig-Thurau et al., 2010; Litvin et al., 2008). The term "social customer" was also recently introduced, depicting the customer who, as a member of online communities, is a free provider of actionable knowledge; that is, knowledge being the result of actions that creates value for the company as it generates new insights into best practices. He is a participant in online conversations as a "word of mouth" influencer, member of customer feedback networks as well as discussion and problem solving communities. The social customer created the need for companies to see the development of strong relationships with customers as a collaborative effort. By engaging them in discussion and activity and by observing, redirecting and even participating in online conversations, customers are seen as partners from the beginning of development and improvement of products and services. Sentiment Analysis, Social Media Monitoring and similar techniques have been developed in order to help companies gain access to and learn about dominant customer insights (Greenberg, 2009). Through virtual communities and social media, customers have the ability to establish themselves as co-creators of value in innovation efforts, both in ideation and design stages as well as prototyping. Companies can actively use customer communities as idea generation sources and are gradually getting more and more familiar with practices of harnessing consumer intelligence from virtual communities for new product development. External consumer networks can be useful for the identification of problematic features and solution suggestions that very often turn to real new products. Also, examples of transparent online communities engaged in new product prototyping, testing and feedback processes already exist. Lead users, with high experience of product usage and willing risk takers play a significant part in this context (Sindhav, 2011).

7 Collaboration with Customers in the Virtual Landscape

7.1 Pre-requisites for Collaboration and Co-creation

Collaborations and knowledge exchange with customers demand fundamental cultural changes in order to view the customer as a potential knowledge source. Another closely related challenge is that of transforming existing intra-company KM and CRM systems to more open systems that facilitate access by customers and allow for extended interactivity. The usage of a wide range of communication and interaction

technologies is necessary to this end (Gibbert et al., 2002). Also, senior management support in communicating and viewing of the customer as a valuable knowledge source and potential partner is considered to be crucial for a company to develop its CKM competence (Rollins et al., 2005).

Sawhney & Prandelli (2000) identify a number of important factors for knowledge co-creation between customers and firms. From the customers` point of view, the most inhibiting factors are the lack of absorptive capacity, lack of trust and lack of motivation. More specifically, absorptive capacity includes the need for customers to enrich their vocabulary for articulating experiences and improving their understanding but also the need to enhance customer connectivity. Trust refers to the avoidance of opportunistic behaviors to gain customer knowledge and the parallel improvement of firm`s reputation and image. Finally, motivation includes the identification of incentives and customer rights. From the company`s point of view, firms must develop the capacities to identify, absorb, share and deploy valuable customer knowledge.

Moreover, Kristensson et al. (2008), suggest that co-creation is a fundamental shift from traditional customization of products approach, and identify a number of key strategies for successful co-creation with customers in the field of new technology-based services development. These include the identification of users` needs depending on their own setting and situations as well as on the various roles they play, the provision of proper analytical tools to enhance the effectiveness of user involvement, the motivation of users by demonstrating the apparent benefits from their involvement and the formation of heterogeneous user groups to ensure idea complementarity and diversity.

7.2 Incentives for Customers` Participation and Involvement

Game theories have been used to describe interactions between buyers and sellers. Relationships between companies and customers that extend in the long run as an ongoing process can be examined through a non-zero game framework, where both actors do better by collaborating than acting alone. While participating in these collaborative relationships, each participant provides something to the other at a lower acquisition or reproduction cost. As a result, value is created. The creation of appropriate and effective communication environments to nurture collaborations of this kind is essential (Gurgul et al., 2002). Also following a cost-benefit logic, user incentives for participating in user innovation communities have been linked to anticipated innovation benefits that exceed respective costs (von Hippel, 2001). Similarly, Nolan (2007), by decomposing the issue of trust development in online business communities, describes that favorable conditions for contribution and participation of the individual occur when information utility (quality information of high practical value) and power of influence exceed the amount of effort required to participate and acquire information. Sidhav (2011) have suggested that contribution of new ideas in online communities are achieved through processes of social validation, information and inspiration, representing strengthening one`s ties with the community and its goals and accessing bits of information that help in filling knowledge gaps and generating new ideas respectively.

Several researchers have addressed issues of motivation for customer participation in online firm-based communities. Brodie et al. (2011) have revealed three major aspects that define the intensity of customer engagement in online business communities, namely the emotional, cognitive and behavioral aspects. Emotions such as gratitude, trust and sense of belonging, the sharing of information and experiences and the amount of online experiences are examples of the specific aspects that characterize the level of complexity of customer engagement states. Hunter and Stockdale (2009) argued that members of business sponsored online communities become involved in them motivated by the mutual exchange of knowledge regarding products and services of their interest. Wiertz and Ruyter (2007) investigated the drivers of knowledge contribution by customers in firm-hosted online communities. Their findings suggest that factors such as online interaction propensity (users` tendency to virtually interact with strangers) plays a significant role on the amount of knowledge contribution. They also found out that customers who exhibit behaviors of commitment and build strong relationships with other community members as well as the collective of the community are more encouraged to share their knowledge.

7.3 A Company to Customer Virtual Interaction Framework

In this section, we propose a theoretical framework on how customer-generated knowledge can be managed and used within a C-Business context (Fig. 2). This framework depicts, in a simple way how customers and companies can collaborate to create knowledge.

According to Zhang (2011), a company`s customer knowledge system should be composed of two major components: A customer knowledge base providing “static” support for customers seeking knowledge and a social software platform, which gives the appropriate “dynamic” communication channels to customers to interact with the firm. Web 2.0 communication and interaction technologies provide the company with the ability to incorporate and use a wide range of applications and techniques such as harnessing collective intelligence, remixing and enhancing web-based data, developing personalization capabilities, exploiting user generated content, and most importantly, manage knowledge co-creation activities. This will allow the firm to maximize user involvement, cooperation, and value co-production (Boselli et al., 2008).

Along the horizontal axis we draw the proposition that customers through virtual networks can more vibrantly develop a mix of competitive and cooperative abilities and actions in terms of knowledge acquisition, sharing and co-creation. In order to intensify knowledge co-creation and cooperation with customers the firm can more effectively use collaboration technologies and stimulate motivation to customers to participate in interactive procedures and share, exchange and create knowledge.

However, forming strategies to achieve more effective knowledge and value co-creating relationships with customers is becoming a less and less straightforward effort as customer usage of social media and the amount of relevant applications and techniques are becoming more massive. According to Zhang (2011) a firm should be able to assess the relative costs of adopting or adjusting social software as well as its effects on existing customer knowledge base index, eventually with the aim of reaching an optimal level of social software.

Also, to our opinion, businesses are also facing the challenge of converging marketing and knowledge management strategies as far as customer knowledge is concerned. Hanna et al. (2011) indicates that customers are gradually participating more actively in the media processes, and adopt a range of participating and value adding roles, such as content creators or critics. This urges businesses to combine traditional marketing promotion and communication activities with further engaging customers in rich interactive dialogues. Similarly, Brodie et al. (2011) suggest that in specific virtual brand communities control is shifting within them as customers are getting more powerful as co-creators of brand meanings, forcing marketers to promote tasks of knowledge sharing and product co-development with customers.

Moreover, firms are forced to consider the complex set of motives and drivers that define customer participation, knowledge contribution and interaction in virtual customer communities. Fuller et al. (2007), in their effort to explore joint innovation activities in online consumer communities, showed that the most active contributors were driven by excitement and pure interest in the innovation activity itself, rather than the plain need of improving a particular product. On the other hand, Wiertz and Ruyter (2007) showed that commitment towards the firm exhibited negative results in terms of quality of knowledge contribution compared to commitment towards the firm-hosted community, which had a positive effect in both quantity and quality of knowledge contribution.

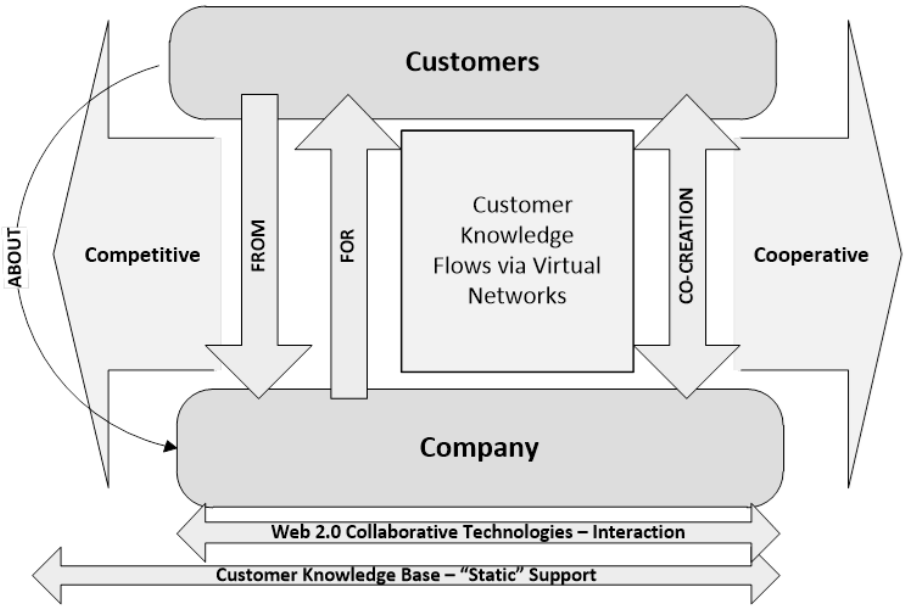


Fig. 2. Company to Customer Virtual Interaction Framework

8 Conclusions

The unsteady and fragile economic and business environment has led firms to collaborate with their stakeholders over the internet in order to create collaborative networks, especially since the rise of Information and Communication Technologies (ICTs) and E-business. Stakeholders within and beyond the value chain cooperate in virtual value networks in order to create and exchange knowledge, and therefore create value for customers.

The growth of Web 2.0 network technologies and of social software is raising customers to a new more powerful level in a company`s virtual value network. In this context, customers` competitive and cooperative abilities have been enhanced. Customers as knowledge actors can be engaged, individually or collectively, in knowledge exchanging and co-creating ongoing relationships of various interaction intensities, depending on specific context, situation and needs. Our framework suggests that Customer Knowledge Management typologies and research can be valuable for companies to develop new strategies that integrate the online building of customer relationships with knowledge sharing and co-creation activities and puts emphasis on the advanced co-opetitive position of customers, that are becoming more active, creative and knowledgeable throughout the virtual landscape.

Additionally, our literature review leads to the conclusion that companies, in order to achieve both task-specific and long-term knowledge co-creative cooperation with customers, should take under consideration a number of mainly intrinsic motivations and drivers that seriously affect customer involvement and contribution. Therefore our further research should elaborate the aforementioned factors, and aim to address the need to strategically manage customers as powerful, co-opetitive virtual players. Zineldin (2004) has proposed that successful, mutually beneficial relationships in a business context have a series of necessary preconditions that include tasks dealing with attitudes, motives, needs, actions and interactions. We have shown that the latter concepts are extensively discussed in the relevant literature concerning knowledge-based relationships with customers, and they have to be further consolidated in a common strategic framework, supported by relevant case studies. Moreover, similar to Fuller et al. (2007) suggestions, our future perspective should also include the strategic challenge that emerges from the distinction between integrating customers for specific knowledge co-creation purposes and building continuous collaborative relationships as constant knowledge sources. This last case is becoming a very complex task as customers` social interactions in the virtual landscape intensify and the amount of unstructured collective knowledge increases (Gruber, 2008).

Acknowledgements. This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II. Investing in knowledge society through the European Social Fund.

References

Afuah, A.: How much do your co-opetitors' capabilities matter in the face of technological change? Strategic Management Journal 21, 387–404 (2000)

Bengtsson, M., Kock, S.: Coopetition in business networks – to cooperate and compete simultaneously. Industrial Marketing Management 29(5), 411–426 (2000)

Boselli, R., Cesarini, M., Mezzanzanica, M.: Customer knowledge and service development, the web 2.0 role in co-production. In: Proceedings of World Academy of Science, Engineering, and Technology, Parigi, vol. 30 (July 2008)

Brodie, R.J., Ilic, A., Juric, B., Hollebeek, L.: Consumer engagement in a virtual brand community: An exploratory analysis. Journal of Business Research (2011)

Buchnowska, D.: Customer Knowledge Management Models: Assessment and Proposal. In: Research in Systems Analysis and Design: Models and Methods, pp. 25–38 (2011)

Chatti, M.A., Klamma, R., Jarke, M., Naeve, A.: The Web 2.0 driven SECI model based learning process. In: Seventh IEEE International Conference on Advanced Learning Technologies, ICALT 2007, pp. 780–782. IEEE (2007)

Consoli, D., Patrucco, P.P.: Complexity and the Coordination of technological Knowledge: the Case of innovation Platforms. Dipartimento di Economia "S. Cognetti de Martiis" LEI & BRICK - Laboratorio di economia dell'innovazione "Franco Momigliano", Bureau of Research in Innovation, Complexity and Knowledge, Collegio Carlo Alberto. WP series, University of Turin (2010),
http://EconPapers.repec.org/RePEc:uto:labeco:201007

Dagnino, G.B., Padula, G.: Co-opetition Strategy: A New Kind of Inter-firm Dynamics for Value Creation. Presented at EURAM: Second Annual Conference – Innovative Research in Management, Stockholm (2002),
http://ecsocman.hse.ru/data/977/644/1219/coopetition.pdf

Eng, T.Y.: Implications of the internet for knowledge creation and dissemination in clusters of hi-tech firms. European Management Journal 22(1), 87–98 (2004)

Fiore, A.M.: The digital consumer valuable partner for product development and production. Clothing and Textiles Research Journal 26(2), 177–190 (2008)

Füller, J., Jawecki, G., Mühlbacher, H.: Innovation creation by online basketball communities. Journal of Business Research 60(1), 60–71 (2007)

Garcia-Murillo, M., Annabi, H.: Customer knowledge management. Journal of the Operational Research Society, 875–884 (2002)

Gebert, H., Geib, M., Kolbe, L., Brenner, W.: Knowledge-enabled customer relationship management: integrating customer relationship management and knowledge management concepts. Journal of Knowledge Management 7(5), 107–123 (2003)

Gibbert, M., Leibold, M., Probst, G.: Five styles of customer knowledge management, and how smart companies use them to create value. European Management Journal 20(5), 459–469 (2002)

Gnyawali, D., Park, B.-J.: Co-opetition and technological innovation in small and medium-sized enterprises: a multilevel conceptual model. Journal of Small Business and Management 47(3), 308–330 (2009)

Greenberg, P.: The impact of CRM 2.0 on customer insight. Journal of Business & Industrial Marketing 25(6), 410–419 (2010)

Gruber, T.: Collective knowledge systems: Where the social web meets the semantic web. Web Semantics: Science, Services and Agents on the World Wide Web 6(1), 4–13 (2008)

Gurgul, G., Rumyantseva, M., Enkel, E.: Customer Integration-Establish a constant bilateral knowledge flow. Institute of Management, University of St. Gallen, HSG (2002)

Hanna, R., Rohm, A., Crittenden, V.L.: We're all connected: The power of the social media ecosystem. Business Horizons 54(3), 265–273 (2011)

Hemetsberger, A., Reinhardt, C.: Sharing and creating knowledge in open-source communities: the case of KDE. In: Paper for Fifth European Conference on Organizational Knowledge, Learning, and Capabilities, Innsbruck (2004)

Hennig-Thurau, T., Malthouse, E., Friege, C., Gensler, S., Lobschat, L., Rangaswamy, A., Skiera, B.: The impact of new media on customer relationships. Journal of Service Research 13(3), 311–330 (2010)

Holsapple, C.W., Singh, M.: The knowledge chain model: Activities for competitiveness. Expert Systems with Applications 20(1), 77–98 (2001)

Hunter, M.G., Stockdale, R.: Taxonomy of online communities: ownership and value propositions. In: 42nd Hawaii International Conference on System Sciences, HICSS 2009, pp. 1–7. IEEE (2009)

Katsanakis, I., Kossyva, D., Vorria, E.: Co-opetition Strategy in Business Excellence: Confronting the Economic Crisis. In: Proceedings of 14th QMOD-ICQSS Conference on Quality and Service Sciences, August 29-31, pp. 949–958 (2011) ISBN: 84-8081-211-7

Katsanakis, I., Kossyva, D.: C-Business: A Theoretical Framework for the implementation of Co-opetition Strategy in E-Business. Procedia - Social and Behavioral Sciences 58, 259–268 (2012)

Kim, S., Zhu, J., Smari, W.W., McQuay, W.K.: Security and Access Control for a Human-centric Collaborative Commerce System. In: Proceedings of the 2006 International Symposium on Collaborative Technologies and Systems, pp. 429–439 (2006)

Koeniq, M.E.D.: KM Moving into Stage IV? The Extra-Organizational Stage. In: Knowledge Management in Practice: Connections and Context, pp. 9–16. Information today, Inc., Medford (2008)

Kossyva, D., Katsanakis, I., Georgopoulos, N., Sfakianakis, M.: Value Creation and Value Appropriation through C-Business. Presented at the 7th National and International Conference of Hellenic Society for Systemic Studies (HSSS): Professional Systemics in Action, Athens, Greece, May 4-7,

Kristensson, P., Matthing, J., Johansson, N.: Key strategies for the successful involvement of customers in the co-creation of new technology-based services. International Journal of Service Industry Management 19(4), 474–491 (2008)

Lawer, C.: On customer knowledge co-creation and dynamic capabilities. Computer and Information Science, Miscellaneous Papers (2005), http://www.mendeley.com/research

Lesser, E., Mundel, D., Wiecha, C.: Managing customer knowledge. Journal of Business Strategy 21(6), 34–37 (2000)

Liberona, D., Ruiz, M., Fuenzalida, D.: Customer Knowledge Management in the Age of Social Networks. In: Uden, L., Herrera, F., Bajo, J., Corchado, J.M. (eds.) 7th International Conference on KMO. AISC, vol. 172, pp. 353–364. Springer, Heidelberg (2013)

Litvin, S.W., Goldsmith, R.E., Pan, B.: Electronic word-of-mouth in hospitality and tourism management. Tourism Management 29(3), 458–468 (2008)

Maier, R.: Modeling knowledge work for the design of knowledge infrastructures. Journal of Universal Computer Science 11(4), 429–451 (2005)

Najmaei, A., Sadeghinejad, Z.: Competitive Strategic Alliances Through Knowledge Value Chain. International Review of Business Research Papers 5(3), 297–310 (2009)

Nolan, T., Brizland, R., Macaulay, L.: Individual trust and development of online business communities. Information Technology & People 20(1), 53–71 (2007)

Nonaka, I., Takeuchi, H.: The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford UniversityPress, Oxford (1995)

Patrucco, P.P.: Collective knowledge production costs and the dynamics of technological systems. Economics of Innovation and New Technology 18(3), 295–310 (2009)

Polanyi, M.: The tacit dimension. University of Chicago Press (2009); first published in: 1966 Routledge and Kegan, London (1967)

Rigby, D., Zook, C.: Open-market innovation. Harvard Business Review 80(10), 80–93 (2002)

Ritala, P., Hurmelinna-Laukkanen, P.: What's in it for me? Creating and appropriating value in innovation-related co-opetition. Technovation 29, 819–828 (2009)

Rollins, M., Halinen, A.: Customer knowledge management competence: towards a theoretical framework. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences, HICSS 2005, p. 240a. IEEE (2005)

Salomann, H., Dous, M., Kolbe, L., Brenner, W.: Rejuvenating Customer Management: How to Make Knowledge For, From and About Customers Work. European Management Journal 23(4), 392–403 (2005)

Sawhney, M., Prandelli, E.: Beyond customer knowledge management: customers as knowledge co-creators, pp. 258–282. Idea Group Publishing, Hershey (2000)

Sigala, M.: E-service quality and Web 2.0: expanding quality models to include customer participation and inter-customer support. The Service Industries Journal 29(10), 1341–1358 (2009)

Sindhav, B.: Co-creation of Value: Creating New Products through Social Media. International Journal of Management Research 2(1), 6–15 (2011)

Smith, H.A., McKeen, J.D.: Developments in Practice XVIII-Customer Knowledge Management: Adding Value for Our Customers. Communications of the Association for Information Systems 16, 755 (2005)

Von Hippel, E.: Learning from open-source software. MIT Sloan Management Review 42(4), 82–86 (2001)

Wagner, S.M., Eggert, A., Lindermann, E.: Creating and appropriating value in collaborative relationships. Journal of Business Research 63, 840–848 (2010)

Wasko, M.M., Faraj, S.: Why should I Share? Examining Social Capital and Knowledge Contribution in electronic Networks of Practice. MIS Quarterly 29(1), 35–57 (2005)

Wiertz, C., Ruyter, K.: Beyond the call of duty: Why customers contribute to firm-hosted commercial online communities. Organization Studies 28(3), 347–376 (2007)

Zahra, S., George, G.: Absorptive Capacity: A Review, Reconceptualization, and Extension. Academy of Management Review 27(2), 185–203 (2002)

Zanjani, M.S., Rouzbehani, R., Dabbagh, H.: Proposing a conceptual model of customer knowledge management: a study of CKM tools in British dotcoms. In: Proceedings of World Academy of Science, Engineering, and Technology, Rome, vol. 28 (April 2008)

Zhang, F., Deng, S.: Research on Knowledge Sharing Model for Collaborative Commerce. In: Proceedings of the 4th International Conference on Wireless Communications, Networking and Mobile Computing, pp. 1–3 (2008)

Zhang, Z.J.: Customer knowledge management and the strategies of social software. Business Process Management Journal 17(1), 82–106 (2011)

Zineldin, M.: Co-opetition: the organisation of the future. Marketing Intelligence & Planning 22(7), 780–790 (2004)

Designing and Developing a Business Continuity Plan Based on Collective Intelligence

Efstathios Fasolis[1], Vassilios Vassalos[2], and Angelika I. Kokkinaki[3]

[1] Information Systems Program, Open University of Cyprus, Cyprus
efstathios.fasolis@st.ouc.ac.cy
[2] Department of Informatics, Athens University of Economics and Business, Greece
vassalos@aueb.gr
[3] Department of Management and MIS, University of Nicosia, Cyprus
kokkinaki.a@unic.ac.cy

Abstract. This paper proposes a methodological approach that supports Collective Intelligence towards the design and development of a Business Continuity Plan (BCP), in order to minimize the potential of a disaster in the organization. In this framework, Collective Intelligence (CI) is supported by Web2.0 technologies that act as a diagnostic tool, providing the ability to the community of an organization to contribute with their collective experience and their intelligence, in the resolution of factors affecting the success of the Business Continuity Plan (BCP). E-BCP, the platform developed for the BCP and it is supported by CI, is also presented. Some open research issues have been outlined.

Keywords: Collective Intelligence (CI), Business Continuity Plan (BCP), Social Networks for optimizing business processes.

1 Introduction

In recent years, a number of risk instantiations of significant importance have affected the business sector and more specifically the business continuity of organizations; some of such incidents have even been detrimental to the viability of the affected organizations. These incidents have pronounced the need for a Business Continuity Plan (BCP), a structured action plan to be followed in cases of adversity, which is considered to add value to organizations, creating competitive advantages, savingmoney, time and resources [Harris (2010)]. A major success factor for a BCP in every organization, are the people who develop it and the organization community that accepts and contributes to its proper implementation. Factors, such as inefficient communication, participation and notification of the plan to the employees of the organization, may also affect the success of the BCP to a great extent.

Web 2.0 technologies and tools have been examined from many different angles, including aspects with respect to consumers and businesses; emerging behaviors; models; ways in which information is searched, evaluated, produced, consumed and services are formulated [Bonabeau (2009)]. In this paper, we examine how the notion

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 278–285, 2013.

of collaborative collective intelligence (CI) could be harnessed to contribute to the formation Business Continuity Plan (BCP).

A major challenge of this research was the ascertainment of minimum references on the design or the development of a Business Continuity Plan (BCP) based on Collective Intelligence (CI). This challenge was at the same time a strong motivation to examine how the Collective Intelligence could support critical aspects affecting professionals in the area of Business Continuity.

The remaining of this paper is structured as follows. Section 2 outlines the Business Continuity Plan, the process for its formation and the factors that affect a BCP. Section 3 outlines the existing platforms for Collective Intelligence. Section 4 examines how a Collective Intelligence platform could be used to support the Business Continuity Plan and section 5 concludes this paper.

2 Business Continuity Plan: Development Process and Critical Success Factors

The methodology for the development of a BCP in this paper is based on the BS25999 [BS25999 (2006)]. The lifecycle of the Business Continuity Management (BCM) consists of six elements and can be implemented by organizations of all sizes and sectors. The effort is adjusted according to the needs of each organization. These elements are described below:

2.1 BCM Program Management

The program management provides the ability to establish and maintain business continuity in a manner appropriate to the size and complexity of the organization [BS25999 (2006)]. This is the business continuity policy which states the objectives of the BCM developed by the organization. The key to this element is gaining top management commitment and assigning the appropriate roles and responsibilities to the BCM program team.

2.2 Understanding the Organization

The aim of this element is to provide important information to assist the understanding of the organization products, assets and services. This is performed through a process which is commonly referred to as a Business Impact Analysis (BIA).The objective of BIA is to determine the acceptable time frame restoration and the allowable data loss, which does not affect the business continuity. The first stage of BIA is to identify and prioritize the critical applications/processes and activities performed through the Risk Assessment (RA). This comprises the ongoing process of identifying threats, for assessment of their risk (in terms of impact and likelihood), and driving activities to avoid such risks or reduce such risks to acceptable levels.

Criticality Assessment. The purpose of a criticality assessment is to establish how critical a system or process is for an organization in the event of loss of confidentiality, integrity, or availability [Harris (2010)] An example is given in Table 1.

Table 1. Criticality Assessment

Dimension	Comments
C (Confidentiality)	Loss of confidentiality, i.e. information may be disclosed to the wrong people.
I (Integrity)	Loss of integrity, i.e. information may be falsified or otherwise corrupted
A (Availability)	Loss of availability, i.e. information may not be available for processing or use for a specific period of: • ½ day • 1 day • 2 days • 1 week • 1+ month

Criticality ratings (i.e. for integrity and availability) and critical timescales for availability must be verified against the Maximum Tolerable Downtime (MTD), that is, the time period between a disruptive event and restart of normal processing.

Classification of Systems and Processes. To assure the criticality classification accuracy of the organization systems/process a benchmark method should be decided. One option could be the creation of a questionnaire evaluating the impact that may cause an organization loss according to the CIA (Confidentiality, Integrity, and Availability) triad and the MTD [Harris (2010)] An example is shown in Table 2.

Table 2. Systems and processes classification versus MTD

Criticality classification	Classification description	Maximum Tolerable Downtime (MTD)[1]
A	Extremely critical	0-24 hours
B	Very highly critical	1-3 days
C	highly critical	up to 1 week
D	Important	between 1-4 weeks
E	Important but not critical	1+ month

2.3 Determining Business Continuity Strategy

As a result of the previous element (inputs and outputs of BIA), an organization will have to evaluate and choose the appropriate continuity strategies in order to meet the

[1] The MTD values depend on the size and complexity of the organization.

organizations business continuity objectives. The selection and evaluation of the applicable business continuity strategy depends on the Business Impact Analysis (BIA). The critical systems/processes have been identified; assessed, classified and critical timescales for availability have been considered according to the MTD.

2.4 Developing and Implementing BCP's

This element focuses in the development and implementation of the plans. These plans should include all the necessary actions required to meet the organizations business continuity objectives in case of a disaster. Individuals or teams with good business knowledge and expertise in certain domains of organizations are required to be involved. Every plan should be practical, complete, efficient, and include step-by-step instructions on what should be done and who will do it.

2.5 Maintaining and Testing the BCP

The aim of this element is to ensure that all the plans are reviewed, updated and tested periodically, to make sure that all arrangements are in place in order to safeguard the organization from a possible disaster. A BCP is a dynamic model that changes as the organization progresses in time.

2.6 Embedding BCP in the Organization's Culture

The final element aims for the BCP to become a valuable part to enhance an organizational BCP culture. The key to this element is communication, participation and notification of the plan to the employees of the organization.

It should be emphasized that an outdated BCP could be insufficient and may provide a company with a false sense of security; this could be devastating if and when a disaster actually takes place. Companies need a live and flexible structured action plan. In light of these, the following important factors affecting the success and the validity of a BCP in an organization [Harris (2010)] ought to be considered:

- Management commitment
- Effective communication, participation and communication of the plan to the employees of the organization
- Proper assignment of roles and responsibilities among the project team
- Proper identification and prioritization of critical applications/processes and resources
- Identification of the risks that the organization needs to be protected
- Test and regular review of the plan, in order to adjust any new developments and changes in the organization
- A motivated environment inspiring innovative ideas.

When addressing business continuity planning, some organizations may focus only on backing up data and on providing redundant hardware. Even though these aspects are

important, these are fragmented attempts towards the organization's overall operational capability. Working and planning a BCP must include the right people at the right places, establish alternative communication channels, and ensure that all processes and applications are properly understood and taken into consideration.

3 Collective Intelligence, Motivation and Platforms

Collective Intelligence (CI) may emerge extemporaneously and could be developed in many forms and setting. In recent years, the extensive use of Web 2.0 technologies have contributed to the development of services (i.e. Amazon Turk, Wikipedia, Google etc.) that encompass the principles of collective knowledge and collective problem solving and motivate a virtual community to participate in an innovative idea or design thinking. Various techniques have been reported [Surowiecki (2004)] for the exploitation of CI in organizations aiming to solicit knowledge and experience from the community, discover innovative ideas, enhance skills in various sectors while at the same time they save money, time and resources. Key to the success of CI is the ability of the organization to identify and highlight the appropriate motives in a community or an organization. Some motivational factors identified in [Malone (2009)], are listed below:

- Ideology – participation in a collaborative intelligence contributing to something that it is believed that will make a difference in society
- Challenge – participation which provides a sense of personal fulfillment or self-realization through the acquisition of additional knowledge and skills
- Career – participation in the development of the individual's career
- Social – desire to exchange experiences with others
- Fun – join for entertainment, enjoyment and rest from other activities
- Reward –participation to receive tangible rewards such as money, gifts and prizes
- Recognition – join in order to receive private or public acknowledgement
- Duty – join in response to a promise or commitment to something.

Those organizations that are able to acquire knowledge and use it in such way that they create their own intelligence and collective solutions become more competitive. During the last years, a number of Collective Intelligence (CI) platforms have been proposed. These CI platforms may be conceptualized into two major categories.The first allows the crowd to collectively contribute to a project; for example, the design of an information system. The second leads the crowd to a simple and distinct problem solving and decision making; for example the "Amazon Mechanical Turk" [http://www.mturk.com. In this platform, challenges or problems are published to an unknown number of users; ideas and solutions are solicited through an open invitation. The users referred as the crowd, form and submit solutions, These solutions are evaluated by the original entity assigned the problem/challenge. Reimbursement may be cash, prizes or recognition.

4 E-BCP: A Platform That Supports Collective Intelligence for a Business Continuity Plan

Collective Intelligence (CI) has been used in organizational settings to achieve increased innovation, productivity and responsiveness, reducing the time required to access the proper knowledge. Could CI assist in the development and design of a BCP by harnessing information, solutions and reducing uncertainly decision making during or before a disaster?

Based on the BS25999 Business Continuity Management (BCM) lifecycle and existing applications of Collective Intelligence [BS25999 (2006)], we reckon that Collective Intelligence can support some BCP elements, as shown in Figure 1.

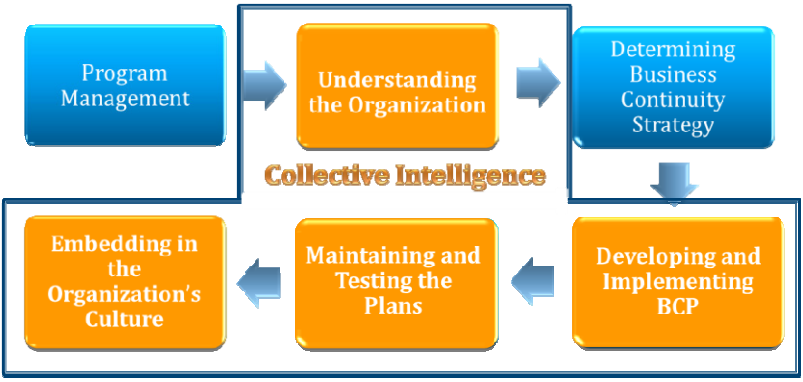


Fig. 1. BCM lifecycle elements which can be supported by CI are highlighted

Fig. 2. A screenshot of the E-BCP platform

As a proof of concept, a collective intelligence platform was designed and developed (www.e-bcp.eu) enabling the users community within an enterpise in Greece to contribute their collective experience and intelligence to resolve factors that affect the success of the BCP. A screenshot of the platform (Greek version) is given in Figure 2.

This service platform was powered by innovation centric collaboration techniques, best practices and processes bundled into a single service conducive for open innovation. Mainly, the platform provides a Web 2.0 collective tool rewarding the knowledge and exploiting the intelligence of the organization's community to implement a BCP. Two tools have been implemented to support the "e-BCP" platform. These tools are shown in Figure 3.

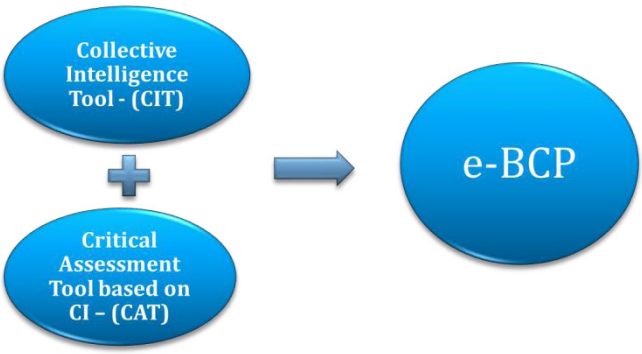


Fig. 3. e-BCP collective Intelligence Platform

The first tool "Collective Intelligence Tool (CIT)" is based on the ability of the organization's BCP project team to raise challenges based on BCP issues that require resolution. Challenges included: the update of manuals and plans; suggestion on how to handle risks and recommendation on how to promote the BCP program within the organization's community. The user community can act as a problem solver and can identify the challenge(s) in the area their expertise and domain of their interest. The user individuals that solved the challenges may get reward, recognition and intellectual satisfaction. Once the predefined schedule for making solution is over, all the solutions created by solvers are submitted to the BCP project seeker team. The BCP project team (seeker) evaluates the solutions, and takes the final call on which solution meets the expectations most optimally and rewards the corresponding solver.

The second tool "Critical Assessment Tool (CAT)" also uses the above CI approach to apply a BIA analysis creating a critical league consisting of all organizations systems/processes. This is achieved by raising several questionnaires to the organizations user community contributing their expertise identifying critical services or spotting any new ones.

Employees became comfortable sharing their experience, thoughts and ideas that can be used to address critical business challenges that may assist the design and the development of a BCP within the specific organization. Professionals involved in business continuity had the ability to filter out the critical information through the

chaos of information which is gathered. Preliminary evidence supports that Collective Intelligence should not be absent from the design of Business Continuity Planning providing a different approach in understanding the organizations business environment harnessing proper information from its community.

5 Conclusions and Future Research

This paper claims that Collective Intelligence is a valuable tool supporting business continuity for harnessing information and resulting to solutions by reducing uncertainly decision making during or before a disaster. Collective Intelligence is an asset, which if collected and used correctly has the potential to emerge as a significant tool to implement and design a BCP providing a strategic advance for an organization such as saving money, time and resources.

The value of this paper is twofold. First, it has provided a proof of concept that this is both feasible and contributes positively in the formation of a BCP in an organization. Second, it sets the direction for further research on the interdisciplinary issues of BCP and CI.

The almost complete lack in relevant bibliography denotes that further research in the theory of Business Continuity based on Collective Intelligence should be prioritized. Extending this study is to investigate the approach utilizing the cooperation and alliance between organizations supported by the logic of collective intelligence regarding business continuity maters. Tools offered by the technologies of "Cloud Computing" can be utilized to this direction. Finally, it would be interesting to consider the opposite approach of Collective Intelligence on the mismanagement of the user community within the organization on issues relating to business continuity.

In conclusion, the field of business continuity supported by the logic of collective intelligence is in the early stage. It has been identified that it is possible to address aspects of BCP using CI. However a number of research topics remain open and need further exploitation.

References

1. BS 25999 – 1 2006: Business Continuity Management – Part 1 Code of practice. British Standards
2. BS25999/BS 25999 News: Dedicated to the Emerging Business Continuity Management Standard, British Standards Institution (2006), http://www.bs25999.net/2006_07_01_archive.html
3. Malone, T., Laubacher, R., Dellarocas, C.: The Collective Intelligence Genome. Journal of MIT Sloan Management Review 51(3) (2010)
4. Surowiecki, J.: The wisdom of the crowds: Why the many are smarter than the few. Random House Inc., New York (2004)
5. Bonabeau, E.: Decisions 2.0: The Power Collective Intelligence. Journal of MIT Sloan Management Review 50(2) (2009)
6. Dierks, E.: Intelligence-Based Business Continuity Part 1, 2, 3. Article of Continuity Insights (2012), http://www.continuityinsights.com
7. Harris, S.: CISSP All In One Exam Guide, 5th edn., pp. 777–840. McGraw-Hill, New York (2010)
8. Amazon Mechanical Turk: Human intelligence tasks, http://www.mturk.com

Rewarding In-Game Banner Ad Clicks with Tangible Incentives

Ruey Komulainen[1,*], Waqar Nadeem[2], Saara Satokangas[1], and Jari Salo[1]

[1] Department of Marketing, University of Oulu
{rukomla,saarasat}@mail.student.oulu.fi,
jari.salo@oulu.fi
[2] Martti Ahtisaari Institute of Global Business and Economics, University of Oulu
waqar.nadeem@oulu.fi

Abstract. Mobile games in the past years have become a new, growing channel for advertisers to reach their customers. The wide spread use of mobile devices, such as smart phones, tablets and PDAs, facilitates the delivery of a variety of highly targeted interactive advertising messages including the use of mobile games as a media. As the phenomenon is new, the literature on mobile in-game advertising is very scarce. The objective of this research is to bridge this gap by studying the consumers' general attitudes towards mobile in-game banner advertising (IGBA) as an advertising format and how effective is reward incentives affects the game players' attitude towards banner advertisements and consequent actions in the mobile game context. A survey research was completed in October 2012, with 426 responses from Finland. The results show that tangible and flexible reward scheme with a link to location based rewards where mobile gamers can redeem physically by converting the points achieved from their game plays, are effective to positively affect their attitudes and intention to take actions.

Keywords: Mobile Advertising, Game Advertising, Ad Clicks, Rewards, Tangible Incentives.

1 Introduction

Mobile technology has increased the possibilities to reach and serve consumers in multiple ways anytime, anywhere relatively easily and cheaply in a highly targeted, interactive way [34], [37]. The increasing popularity of mobile gaming [28] has enhanced the appeal of using mobile games as a marketing medium which requires more attention from researchers as well as practitioners. Growth in mobile gamers was 35% year-on-year in US totaling 100 million, whilst in EU the growth was 15% year on year totaling 70 million calculated based on five countries including Finland in 2012 [28]. Engaging nature of games, and long exposure to the advertisements are seen as major benefits of advertising through games compared to traditional media

* Corresponding author.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 286–297, 2013.

[8],[27],[36],[41]. According to [28], gaming took 13% of all time spent on games worldwide, totaling more than 130 million hours a day and 9% total money spent on games, grossing USD 5.8 billion. According to [29], the usage of mobile web has been growing massively in the past years. Almost half of under 45-year-olds use internet weekly on their mobile phones outside their homes [32]. In short, mobile games offer a pertinent way to reach and engage with consumers, especially young adults [31].

The research in mobile marketing itself is still in its early stages [34], and there is even less research on mobile games as an advertising medium. The literature on mobile in-game advertising is at embryonic stage. Hence, the literature on mobile in-game advertising effectiveness is very scarce although there are researches on mobile advertising effectiveness in other formats such as text messages [39], [2]. As such, this research aims to close part of this observed gap by studying the consumers' attitudes towards mobile in-game banner advertising as an advertising format and how effective is location-based reward incentives affects the game players' attitude towards banner advertisements and what are their intentions to take action in the mobile game context.

This research reviews existing literature which is used to build the theoretical framework. Afterwards, methodology and the empirical analysis are presented. Finally, the results are discussed, including the limitations of the existing research along with discussion and conclusion

2 Literature Review

In-games advertising (IGA) refers to the incorporation of brands into existing games [43]. Mobile advertising itself, is not a new innovation but the formats of doing mobile advertising keep evolving at a fast pace. Before the wide adoption of smart phones, PDAs and tablets, SMS and MMS messaging were the primary advertising formats in the mobile medium [2], [34], [42]. Much of the research on acceptance and attitudes towards mobile marketing is made based on the basic mobile phone [34]. Previous studies look at acceptance of different messaging formats, when the advertiser sends either a solicited or unsolicited message to a consumer [2], [11], [16], [39], [42]. However, today high-tech mobile devices along with new mobile technologies provide increasingly multisided medium with various m-commerce, marketing and advertising possibilities [19], [34]. There is now a wide range of formats that are unique to the mobile medium such as location-based advertising, display advertising inside different mobile applications and mobile coupons [34]. This is why further research is needed on how consumers perceive them.

2.1 Attitudes towards Advertising in Different Media

As banner advertisements are widely used in internet advertising, there is quite significant amount of research looking at the effectiveness and consumers' attitudes towards them. [4] states that consumers perceived banner advertisements more

favorably than other online ad formats because they perceived banners as less intrusive ad formats. On the contrary, [10] identifies past negative experience with banners causes people to avoid the source in the future. The phenomenon is sometimes referred as "banner blindness" [10]. Also, [35] states that consumers are often engaged in selective perception resulting in limited information exposure. However, in mobile games context, the screen size is relatively small and the game players are engaged in gaming activities, thusly, the attitude towards banner format is unclear.

Due to the advanced technology and wide adoption of mobile devices, mobile medium provides new possibilities for delivering up-to-date, relevant, customized data easily and inexpensively. However, this does not mean that consumers want to receive it. Mobile devices are seen as highly personal and intimate devices, and advertisements send without permission are easily seen intrusive and annoying, [12], [17], [34], [39]. When permission is given, the consumer attitudes are more positive than with advertising without permission [1], [3], [39]. In this research, the permission is acquired when the gamer plays the game. However, a gamer's primary concern is to 'play a game', not to receive advertisements. Similarly, consumers involuntarily face banner advertisements when they go on a website to perform other tasks [9].

Moreover, in-game advertisements are seen as positive, but these positive attitudes are generated only when some relatively strict conditions are filled/ satisfied. According to earlier studies, the in-game advertisements usually have to be congruent with the game theme to be seen positively [18], [23], [38]. [23] states that incongruent advertisements have been perceived as "odd", "surprising", "inappropriate" and "confusing". Even if these conditions are met, the advertisements are not been perceived positively by everyone. Also, game players' comments noted that the world is already filled with commercials and hoped that their game-world would be left alone [27]. Due to banner advertisements rather disconnected nature, they thus might be seen negatively. As such, we hypothesized that:

H1: *Attitudes of game players towards banner advertisements in mobile games are negative.*

2.2 Incentives

For the purpose of this research, mobile coupons refer to digital coupons on a mobile device such as a mobile phone, smart phone, or personal digital assistant. People carry their mobile phones almost always with them, along with the mobile coupons [12]. This medium also reduces costs to businesses related to database creation and maintenance, printing, physical distribution, measurement and tracking. Digital coupons have amplified outreach allowing businesses to obtain immediate quantitative information, thus enable to adapt target their consumers' preferences quickly [20]. Promotional offers, such as discount coupons, free samples or lucky draws, have been used to enhance consumer attitudes towards mobile marketing, and to increase the effectiveness and acceptance of mobile marketing [1], [3], [11], [39]. Thus, they could be used to enhance the effectiveness of mobile in-game banner advertising as well.

Interactive properties of games can also be applied to game playing such that gamers can make choices or get rewarded through involvement with the brand [22]. [17]

studies rewards based on gamer's participation in the game play as well as achievements in the game. They found that when consumers are voluntarily connected to the campaign via the offered prizes, it will lead to more responses and more affective attitude toward the brand and the company. Similarly, the positive feeling of challenge and accomplishing a goal may be transferrable to a real life brand. Rewarding consumers for their achievements is not only beneficial in terms of increased redeeming of company's offers, but it might translate into increased amount of playing. The overall high interest in mobile coupons, the even higher interest in mobile coupons as rewards in the game and the extending exposure time to advertisements; all are advantages of this new promotion format [7], [21], [24], [33], [41].

2.3 Factors Affecting Effectiveness of Incentives

2.3.1 Location Based Incentives

Location-based mobile services are services in which the location of a person or an object is used to determine the application or service [3]. Location-based mobile services include mobile marketing based on consumer's location [40]. Commercial and promotional messages are not the only thing that can be tied to the location of the mobile device but also the game playing. Integrating player's location into the game experience is one factor that makes mobile games special compared to traditional video games [6]. Similarly, integrating player's location into the in-game advertising could make mobile IGA special compared to traditional IGA as well.

However, the very idea that game players can be tracked creates sense of loss of privacy [40]. In Finland, the consumers seem to somewhat perceive the benefits greater than the possible negative effects. According to a study reported by [19], in the EU area, 48% of smart phone users are interested in location-based vouchers. And the same number for Finland is 57%. This shows that there is an interest in new marketing innovations such as location-based promotions which has been enabled by advanced technology.

2.3.2 Redemption Options for Mobile Coupons

There are few options for companies to organize the actual redemption of the incentives offered within mobile games. The first one presented here is when consumers store the coupon on their mobile device and show it at the cashier to the store clerk [26]. The other ways could be through e-commerce or m-commerce. It is important to know how consumers perceive these different options, so that marketers can plan their mobile incentive redemption taking consumer preferences into account.

[12] study shows that redemption effort is the main determinant of attitude towards mobile coupons where consumers are concerned about the complication and cost to redeem such coupons. When looking at mobile games, the players are familiar with the technology. Thus the fear of technological difficulty might not be as severe in this context.

Redemption can also be done via online computer or mobile webstore. According to a report issued in November 2012 by [32] – already two thirds of the Finnish

people had purchased something online during the past year and the number has been steadily rising. Consumers are thus already relatively comfortable buying online. Again, young adults are the main age group who has adopted this behavior [15]. However, there are no conclusive answers from earlier studies in mobile context, thus our next set of hypotheses are:

H2a: *Tangible rewards in physical world affect mobile gamers' attitude towards mobile IGBAs.*

H2b: *Flexible reward schemes affect mobile gamers' attitude towards mobile IGBAs.*

H2c: *Location based rewards affect mobile gamers' attitude towards mobile IGBAs.*

3 Research Framework

Based on the existing literature about attitudes toward advertising and consumer behavior models, a research framework is constructed using the Theory of Reasoned Action (TRA) proposed by Fishbein and Ajzen in early 1970s. There are three major constructs – Attitude, intention and behavior. The model describes the relations between attitude and behavior by linking individual beliefs, attitudes, intentions and behavior. Attitude toward the behavior is defined as the individual's positive or negative feelings about performing a behavior. It is an assessment of one's beliefs regarding the consequences arising from a behavior [14]. This model has been used quite widely in previous studies to explain the advertising effectiveness for other formats [39], [13], [1], [11].

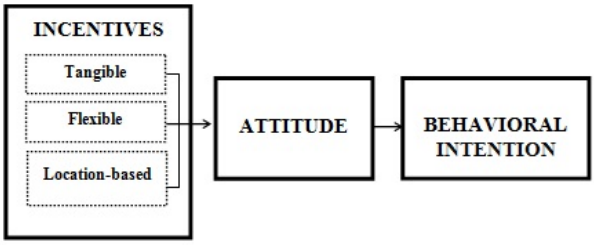


Fig. 1. Conceptual Framework: Adapted from [14] and [39]

4 Methodology

This research is conducted with quantitative methods, which is typical for a deductive approach [5] employed here. The data gathering in quantitative social studies is done from a group people who represent a certain population via a questionnaire designed for the purpose [5]. Questionnaires are used to ensure an orderly and structured approach to data gathering [25]. Thus, in this research the data were collected using an online questionnaire and analyzed with SPSS programs.

Attitudes and intentions towards advertisements have already been widely researched, the survey questions in this area were formulated based on the earlier research. There are valid measures for attitudes and intentions in both online and mobile advertising [39], [13], [1], [11]. The questions regarding mobile game specific – factors and features of incentive systems were self-formulated basis on the earlier research [39], [38] and [12] as well as the interview data.

The survey consisted of a total of 18 questions for those who play mobile games and who had noticed banner advertising in them. A screenshot of a banner ad in mobile game was displayed when asked if the respondents notice similar type of ad when they play mobile games to ensure the respondents' full apprehension of research topic. The questions were mainly closed multiple choice questions but also multiple choice questions with an optional open-ended space were provided to gain more explicit information on how consumers view this new advertising medium [5]. The questions were measured on a Likert scale from 1 to 7 where 1 as 'Strongly Disagree' and 7 as 'Strongly Agree'

Of the different sampling techniques, convenience sampling was incorporated to gather the data for this study. This method is very common in the field of business and management and more prominent than samples gathered based on probability sampling [5]. Although convenience sampling was employed in this research, some procedures were done to ensure the representativeness of the sample and thus reliability of the results. Firstly, the questionnaire was sent mainly to university students, who are mainly of the right age group considering mobile game playing. They also have relatively high skills in English in consideration that the electronic survey is written in English. Secondly, the questionnaire was run by thesis supervisors, co-researchers and piloted with few university students to ensure that respondents understood the questions correctly. This resulted in revision of some parts of the language used. Thirdly, the questionnaire included watershed questions regarding game playing and noticing of the advertisements, to get the opinions of those who play mobile games as well as see the advertisements. In a similar manner, those who had never seen advertisements in mobile games were directed to the questions regarding attitude and intentions if they were to see one. Therefore the responses used in the analysis are from people who play mobile games and have encountered advertisements in this medium.

The survey was conducted in October 2012 through researchers' Facebook pages, Master's thesis group on the internet, via e-mails to all student members of Finanssi, the history, geography and language students in the University of Oulu, students in Oulun Lyseo high school; students from Finnish degree programs and staff in the Kajaani University of Applied Sciences. An incentive was used to increase the response rate: lucky draw of gift cards to S-group stores (one of the biggest retailers in Finland). The survey generated 497 responses in total. After removing incomplete responses and excluding responses from non-Finnish origins, the final sample consists of 426 unique responses.

5 Data Analysis and Findings

5.1 Reliability Analysis

Before reporting findings, we present here the evidence that the data collected are of acceptable quality. Table 1 displays the reliability statistics of each construct. These statistics reveal that Cronbach's alpha of all the variables is over 60%. There are total 26 items and alpha for overall scale is 0.700. This value meets the threshold of 0.700, indicating the acceptable reliability of these scales [30].

Table 1. Reliability of Instruments

Variable	Items	Cronbach Alpha
Tangible Rewards	2	0.749
Flexible Rewards	3	0.832
Location Based Rewards	2	0.741
Attitude towards IGBAs	17	0.609
Intention	2	0.618
Total	26	

5.2 Demographics of Finnish Mobile Gamers

In this survey, there are 63% male respondents and 37% female respondents. 58.4% of the male respondents and 72% of the female respondents play mobile games. This suggests that there may be more females than males playing mobile games. 75.8% of the respondents are 20 to 29 years of age, which is the right age group for our research context. 61.5% of our respondents are earning between 201 to 1000 euros. Unlike earlier research [38] where the gamers spend 2-9 hours a day playing video games, our respondents in mobile games play short sessions. Majority of them (76.7%) plays between 1-15 minutes per session and the main reason for playing mobile game is to kill time while waiting for something to happen.

Table 2. Impact of Incentives on Attitude

Variables	Regression Coefficients	T – Statistics	Standard Errors	P Values
(Constant)	2.777	19.743	.141	.000
Incentives	.262	3.456	.030	.001
R – Square	.069			
F – Statistics	11.941			
Overall P Value	.001			

Table 2 provides the summary of regression analysis of the rewards affecting the attitude of the respondents to mobile IGBA's. The results exhibit the substantial contribution of independent variable to the dependent variable (R – Square = .069 and F – value = 11.941). In addition to this the model is also significant ($P < 0.05$). Individually, the impact of each variable, on dependent variable, is significant ($P < 0.05$).

Hypothesis 2, 3 and 4 states the positive influence of rewards on the attitude. It is also evident from the overall impact of rewards/ incentives (including Tangible, Flexible and Location Based Rewards) results as it is having positive influence (0.262) on the dependent variable (Attitude). This means, individuals who are more interested in rewards tend to show a positive attitude towards the IGBAs.

Table 3. Impact of Flexible, Location and tangible incentive on Attitude

Independent variable	R^2	SE of the estimate	Beta	T	Sig.
Tangible	0.079	0.026	0.280	3.716	0.00
Flexible	0.033	0.027	0.181	2.360	0.019
Location	0.062	0.026	0.250	3.302	0.010

As seen from Table 3, the individual variables tangible rewards, flexible rewards, and location-based rewards show that their impact tends to be positive on the attitude of the mobile gamers to receive IGBA's having regression coefficient 0.079, 0.033 and 0.062 respectively.

Table 4. Impact of Attitude on Intentions

Variables	Regression Coefficients	T – Statis-tics	Standard Errors	P Values
(Constant)	.412	.917	.449	.360
Attitude	.290	4.290	.137	.000
R – Square	.084			
F – Statistics	18.403			
Overall P Value	.000			

In table 4 our hypothesis in present study hypothesized the impact of attitude on the intentions of the gamers as consumers. It is also evident form the regression analysis of the present study that attitude affects the intention positively (0.290) and significantly.

Table 5 portrays the descriptive statistic of the seventeen dimensions of Attitude towards In-Game Advertising in terms of mean and standard deviation with a view to answer Hypothesis 1 which states that game players attitude toward banner advertisements in mobile games are generally negative. The findings show that most influencing factor of the negative attitude towards in-game banner advertising (IGBAs) tends to be with the mean of 5.54 "IGBAs are irritating in games that require deep concentration" followed by "I feel IGBAs are irritating". Specifically the results in the Table 5 demonstrate that our hypothesis that suggests that the mean values of majority of the questions asked for the attitude reveal that, the attitude is generally negative towards IGBAs. The IGBA doesn't seem to be quite entreating for the respondents with the mean value of 1.92. Almost, all the means signify a value less than 5 except the reverse coded questions; which reveal that the attitude towards the ad is generally negative. Overall results suggest that mainstream of respondents are irritated with the IGBA.

Table 5. Items of Attitude

Items of Attitude	Mean	Std. Deviation
I feel that seeing IGBAs is entertaining	1.92	1.219
I feel that IGBAs are pleasant	1.88	1.130
I feel that IGBAs are a good source for timely information	2.40	1.396
I feel IGBAs are irritating*	5.42	1.625
Contents in IGBAs are annoying*	5.00	1.571
I use IGBAs as a reference for purchasing	1.95	1.335
I trust IGBAs	2.26	1.376
IGBAs provide information need	1.99	1.220
IGBAs have no effect on how games are played	3.83	1.820
IGBAs don't interrupt the excitement	3.17	1.789
Overall, I like IGBAs	2.04	1.350
It is acceptable to have IGBAs in free games*	5.22	1.429
It is acceptable to have IGBAs in games that cost 3 euros or less	2.79	1.625
It is acceptable to have IGBAs in games that cost 3 euros or more	1.98	1.344
IGBAs are irritating in games that require deep concentration*	5.54	1.630
IGBAs are irritating in simple games	4.54	1.722
I would be interested in to receive IGBAs based on my location	2.94	1.766

* most influencing factors of the gamer's attitude towards in-game banner advertising.

6 Conclusion and Discussion

The main aim of this research is to study the consumers' attitudes towards mobile in-game banner advertising as an advertising format, and whether reward incentives affects the game players' attitude towards banner advertisements (IGBAs) and consequent action in the mobile game context. Our research questions are answered via four-folded ways:

1. The attitude of game players to IGBAs in mobile game is observed to be generally negative
2. The most influencing negative factors being irritations and annoyances caused by IGBAs on mobile games that require deep concentration.
3. Incentives in form of tangible, flexible and location based rewards have positive and significant impact on gamers' attitude towards IGBAs
4. Positive attitude towards IGBAs increase their behavioral intention to click or redeem the rewards.

The research also confirms that there is positive influence of rewards on the attitude as evident from the results showing a positive influence (0.262) on the dependent variable (Attitude). This means, individuals who are more interested in rewards tend to show a positive attitude towards the IGBAs. The research results also show that attitude affects the intention positively (0.290) and significantly. Also, the results

show that tangible and flexible reward scheme with a link to location based rewards where mobile gamers can redeem physically by converting the points achieved from their game play, is effective to positively affect their attitudes and intention to take actions.

Theoretically, this study deepens and enriches our understanding of Finnish mobile game players' attitude towards IGBAs and the consequent actions. Practically, our research provides valuable suggestions to marketers and business managers to help them reduce uncertainty when building a marketing plan targeting Finnish mobile gamers. This research helps them gain more knowledge about their demographic profiles as a consumer - young adults earning some amount of income and they play mobile games to kill time. It is evident from our analysis that the mobile gamers' attitude towards IGBAs is negative. However, there is positive impact of incentives to their behavioral intention that ultimately increase the likelihood to click on IGBAs. The individual variables - tangible, flexible and location-based rewards have shown positive influence to the gamers' attitude towards IGBAs thus increasing their behavioral intention positively and significantly. In other words, they are motivated by mobile coupons with a flexible scheme where they can use in stores nearby. The results from this study are in accordance to previous researches regarding attitudes of consumers to advertising [39], [10], [16] and the factors affecting effectiveness of incentives [40], [3], [12].

While the results have contributed to the body of knowledge about mobile gamers in Finland, we are unable to discuss similar implications to non-Finnish mobile gamers. Consequently, we propose that other researchers should undertake similar study in other countries. Secondly, the research results point to the positive impact of flexible, tangible and location-based rewards to mobile gamers' attitude towards IGBAs, thus it is important that future research follow through deeper understanding of the type of consumer products and services that are most suited for such reward schemes.

References

1. Altuna, O.K., Konuk, F.A.: Understanding Consumer Attitudes Toward Mobile Advertising and Its Impact on Consumers' Behavioral Intentions: A Cross-Market Comparison of Unites States and Turkish Consumers. International Journal of Mobile Marketing 2(4), 43–51 (2009)
2. Barnes, S.J., Scornavacca, E.: Mobile marketing: the role of permission and acceptance. International Journal of Mobile Communication 2(2), 128–139 (2004)
3. Barutçu, S.: Attitudes towards mobile marketing tools: A study of Turkish consumers. Journal of Targeting, Measurement and Analysis for Marketing 16(1), 26–38 (2007)
4. Burns, K.S., Lutz, R.J.: The function of format: Consumer Responses to Six On-line Advertising Formats. Journal of Advertising 35(1), 53–63 (2006)
5. Bryman, A., Bell, E.: Business Research Methods, 2nd edn. Oxford University Press, Oxford (2007)
6. Casual Games Association: Mobile Gaming: Casual Games Sector Report (2012a), http://issuu.com/casualconnect/docs/winter2012report (cited on September 25, 2012)

7. Chang, Y., Yan, J., Zhang, J., Luo, J.: Online In-game Advertising Effect: Examining the Influence of a Match Between Games and Advertising. Journal of Interactive Advertising 11(1), 63–73 (2010)
8. Chaney, I.M., Lin, K.H., Chaney, J.: The Effect of Billboards within the Gaming Environment. Journal of Interactive Advertising 1(5), 37–45 (2004)
9. Cho, C.H.: The Effectiveness of Banner Advertisements: Involvement and Click-through. Journalism & Mass Communication Quarterly 80(3), 623–645 (2003)
10. Cho, C.H., Cheon, H.J.: Why Do People Avoid Advertising on the Internet? Journal of Advertising 33(4), 89–97 (2004); Choi, Y.K., Hwang, J.S., McMillan, S.J.: Gearing Up for Mobile Advertising: A Cross-Cultural Examination of Key Factors That Drive Mobile Messages Home to Consumers. Psychology & Marketing 25(8), 756–768 (2008)
11. Choi, J., Seol, H., Lee, S., Cho, H., Park, Y.: Customer Satisfaction Factors of Mobile Commerce in Korea. Internet Research 18(3), 313–335 (2008)
12. Dickinger, A., Kleijnen, M.: Coupons going wireless: Determinants of consumer intentions to redeem mobile coupons. Journal of Interactive Marketing 22(3), 23 (2008)
13. Ducoffe, R.H.: Advertising value and advertising on the Web. Journal of Advertising Research 36(5), 21–35 (1996)
14. Fishbein, M., Azjen, I.: Belief, attitude, intention and behavior: An introduction to theory and research. Addison-Wesley, Reading (1975)
15. Flurry Analytics: Mobile Freemium Games: Gen Y Plays, but X Pays (2011), http://blog.flurry.com/?Tag=Mobile+Gaming (cited on October 1, 2012)
16. Gao, Q., Rau, P.P., Salvendy, G.: Perception of Interactivity: Affects of Four Key Variables in Mobile Advertising. Int. J. Hum.-Comput. Interact. 25(6), 479–505 (2006)
17. Gao, T., Sultan, F., Rohm, A.: Factors Influencing Chinese Youth Consumers' Acceptance of Mobile Marketing. Journal of Consumer Marketing 27(7), 574–583 (2010)
18. Huang, J., Yang, T.: The Effectiveness of In-Game Advertising: The Impacts of Ad Type and Game/Ad Relevance. International Journal of Business Management 10(1), 61–72 (2012)
19. IAB: Global Mobile Advertising Market Valued at $5,3 Billion (EUR 3,8 Billion) in 2011 (2012), http://www.iab.net/about_the_iab/recent_press_releases/press_release_archive/press_release/pr-060612_global (cited on November 12, 2012)
20. Jung, K., Lee, B.Y.: Online vs. Offline Coupon Redemption Behaviors. The International Business & Economics Research Journal 9(12), 23–36 (2010)
21. Kiip Inc.: Company web pages (2012), http://www.kiip.me/ (cited on October 14, 2012)
22. Kureshi, S., Sood, V.: Indian gamers' recall, recognition and perceptions of in-game placements. Journal of Indian Business Research 1(4), 252–268 (2009)
23. Lee, M., Faber, R.: Effects of Product Placement in On-Line Games on Brand Memory: A Perspective of the Limited-Capacity Model of Attention. Journal of Advertising 36(4), 75–90 (2007)
24. Marketing WeeklyNews: Kiip; New Consumer-Facing Kiip App Makes Storing and Redeeming Kiip Rewards Easier Than Ever. Marketing Weekly News, 77 (August 11, 2012) http://search.proquest.com.pc124152.oulu.fi:8080/abicomplete/docview/1030725520/13A6F7089126F930706/4?accountid=13031 (cited on November 15, 2012)
25. McDaniel, C., Gates, R.: Marketing Research, 7th edn., p. 35. John Wiley & Sons, Inc., Hoboken (2006)

26. Mobile Marketing Association: Mobile Coupons Gaining Wider Acceptance Across Age Groups and Geography (Cellfire) (2011), http://www.mmaglobal.com/studies/mobile-coupons-gaining-wider-acceptance-across-age-groups-and-geography-cellfire (cited on October 7, 2012)
27. Nelson, M., Keum, H., Yaros, R.: Advertainment of Adcreep Game Players' Attitudes Toward Advertising and Product Placements in Compuet Games. Journal of Interactive Advertising 5(1), 3–21 (2004)
28. Newzoo: Mobile Games Trend Report (2012), http://www.newzoo.com/trend-reports/mobile-games-trend-report/ (cited on September 28, 2012)
29. Nielsen: State of the Media: U.S. Digital Consumer Report, Q3-Q4 2011 (2012), http://www.nielsen.com/us/en/insights/reports-downloads/2012/us-digital-consumer-report.html (cited on October 28, 2012)
30. Nunally, J.: Psychometric theory, 2nd edn. McGraw Hill (1978)
31. Official Statistics of Finland: Televisio ylitti miljardin euron rajan – ennakkotiedot joukkoviestintämarkkinoista 2011. Statistics Finland, Helsinki (2012a)
32. Official Statistics of Finland: Two-thirds of Finns are online shoppers. Statistics Finland, Helsinki (2012b)
33. P4RC Inc.: Company web pages (2012), http://www.p4rc.com/index.html (cited on October 14, 2012)
34. Persaud, A., Azhar, I.: Innovative mobile marketing via smartphones: Are consumers ready? Marketing Intelligence & Planning 30(4), 418–443 (2012)
35. Rosenkrans, G.: Maximizing User Interactivity through Banner Ad Design. Journal of Promotion Management 16(3), 265–287 (2010)
36. Salo, J., Karjaluoto, H.: Mobile Games as an Advertising Medium: Towards a New Research Agenda. Innovative Marketing 3(1), 71–82, 100 (2007)
37. Salo, J.: Digital Advertising in Mobile Games: A Conceptualization and Description of Mobile In-Game Advertising. International Business & Economics Review 1(1), 7–15 (2009)
38. Tangmanee, C., Rustanavibul, J.: Product Placement in Online Games: An Exploration into Thai Gamers' Attitude. Journal of American Academy of Business 17(2), 92–98 (2012)
39. Tsang, M.M., Shu-Chun, H., Ting-Peng, L.: Consumer Attitudes Toward Mobile Advertising: An Empirical Study. International Journal of Electronic Commerce 8(3), 65–78 (2004)
40. Unni, R., Harmon, R.: Perceived Effectiveness of Push vs. Pull Mobile Location-Based Advertising. Journal of Interactive Advertising 7(2), 1 (2007)
41. Wise, K., Bolls, P.D., Kim, H., Venkataraman, A., Meyer, R.: Enjoyment of Advergames and Brand Attitudes: The Impact of Thematic Relevance. Journal of Interactive Advertising 9(1), 27–36 (2008)
42. Xu, D.J.: The Influence of Personalization in Affecting Consumer Attitudes Toward Mobile Advertising in China. The Journal of Computer Information Systems 47(2), 9–19 (2006/2007)
43. Yang, M., Roskos-Ewoldsen, D., Dinu, L., Arpan, L.: The Effectiveness of "in-Game" Advertising: Comparing College Students' Explicit and Implicit Memory for Brand Names. Journal of Advertising 35(4), 143–152 (2006)

Adaptive and Compliant Policy Implementation: Creating Administrative Processes Using Semantic Web Services and Business Rules

Yiwei Gong and Marijn Janssen

Faculty of Technology, Policy and Management, Delft University of Technology,
Jaffalaan 5, 2628 BX Delft, The Netherlands
{Y.Gong,M.F.W.H.A.Janssen}@tudelft.nl

Abstract. Adapting to and complying with frequently changing policy against low costs requires flexible business processes. Yet, existing systems are unrelated, consist of operating silos, involve many human translation tasks and policies are hardcoded in business processes. From a technology perspective, both Semantic Web Service (SWS) and Business Rule (BR) can be used to create flexible business processes that can adapt to changes in policy. Despite their potential these technologies are seldom integrated or used in administrative organizations. SWS can be adopted to present functionality needed in processes, whereas the elicitation of BR is an important ingredient for process creation. In this paper, we derive and test guidelines for creating administrative processes that are dynamical created using SWSs and BRs that are derived from policies. A case study is used to demonstrate and evaluate the guidelines. The guidelines can serve as a foundation for developing agile and compliant architectures.

Keywords: Business Process Management, Business Rule, Semantic Web Service, Compliance, Administrative organization, Flexibility, Adaptability.

1 Introduction

Policy changes frequently and needs to be implemented by administrative organizations against low costs and within a short time frame. This has already resulted in a broad acceptance of Service-Oriented Architecture (SOA), which are merely used for creating static processes [1], but which help to reuse components during design time. As such, process design has been limited to pre-defining processes. Yet, frequent changes in policy require that defining processes should not be labor-intensive and time-consuming in advance but that they are created on the fly during run time. In the effort of achieving higher process flexibility, both Semantic Web Service and Business Rule technologies are used. But they are hardly used in integration, whereas the integration might result in higher levels of adaptability.

Semantic Web Service (SWS) provides a declarative, ontological framework for describing services in a machine-readable format that can also facilitate logical reasoning [2]. Using SWS, service descriptions can be interpreted based on their

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 298–310, 2013.

meanings and can be composed into complex processes corresponding to the needs of a service request [3]. While transforming the organization into a service-oriented manner, the impact and requirement of a SOA on the organization and the way how resources are properly allocated in such a service-oriented organization is still not clear. However, the policy implementation strongly influences the operational processes in public organizations and complicates their use of SWS technology.

A *Business Rule* (BR) is a statement that defines or constrains some aspects of the business. It is intended to assert business structure or to control or influence the behavior of the business [4]. Although there are many knowledge representation techniques that can be used to express BRs, the rule technique is most frequently used in expressing BRs. In a classical rule-based approach, the logic of a process is represented by a set of rules, which are associated with business activities and specify the properties of them [5], e.g. the preconditions of its execution. It is a challenge for the administrative organizations how they can consistently and conveniently use BRs among those different formats.

BR is a common approach in knowledge intensive organizations [6, 7]. Aligning SWS and BR technology allows dynamic creation of processes in which predefinition of processes is no longer necessary. To enable this, administrative organizations must integrate SWS and BR components. Integrating SWS and BR requires the organization to have a proper technical architecture solution, which has been reported in [8]. In this paper, we abstract from the architecture by deriving guidelines to create adaptive processes using both SWSs and BRs, taking into account the flexibility and compliance of the processes. The guidelines are generic, and can be implemented using different technologies. The rest of this article is organized as follows. In the next section we discuss the typical policy implementation procedure in administrative organizations. Then, we discuss the guidelines for administrative organizations to implement policy into operational processes using SWSs and BRs. Thereafter, we use an illustrative case study to demonstrate the modeling process. The paper ends with a conclusion and prospects for future work.

2 Policy Implementation

Roughly speaking, policy implementation takes the new or updated legislation of policy as a starting point, translates this into executable policy and finally implements this in the operational administrative processes [9]. Policy regulates human behavior in the real world [10] and is often written in natural language. Administrative organizations take the policy as input and implement it in their operational processes. For this purpose they need to first translate the policy into a computer-readable format which requires policy-interpretation and the collection of additional information. Once the policy is specified, it can be executed using business processes. Table 1 summarizes the expression of policy at different levels, and their possible formats and users. Policy-making is a political activity dominated by politicians and lawyers which provides often the starting point for implementing policy, whereas policy representation is the knowledge representation of policy that can be understood by computer applications. Operational processes concerns the execution of policy by running business processes.

Table 1. Characterization of policy at different levels

Level	Typical Formats	Typical Users
1. Policy-making	Text	Humans
2. Policy Representation	Ontology models, RIF rules	Inference engine, process engine
3. Administrative pro-cesses	ECA rules, Production rules, Programming languages	Software agent, Web Service, process engine

2.1 Policy-Making

Policy is formulated with the objectives of realizing desired behavior. Policy can be ambiguous and different sources of policy might be conflicting. Therefore, policies are not directly executable and need to be interpreted using extra information. Legal experts need to first determine what the requirements are in the decision-making related to legislation and then find out what materials (documents or other forms of information) should be provided to prove whether the requirements have been satisfied. After that, the providers of those materials and the requirements of those materials all need to be clarified. In the example of the Highly Skilled Migrant policy (see Section 4.1), legal experts need to first identify the income requirement, education level requirement, etc. Then they need to interpret what documents are required, including, for example, an employment contract, salary sheet and diploma. This step might add extra requirements to those of legislation, e.g., "the applicant must provide salary statements of the past 3 months". All these requirements are part of the knowledge necessary to make a decision about the application. After that, they also need to indicate the sources of the required materials or information. For instance, the salary sheet is from the migrant, and the employment contract is from the employer. Different options in collecting certain information are often feasible. Finally, a sequence of document checking and decision making is formed. The policy interpretation can be considered as a high level process design. Although such an interpretation is human based and not understandable by computers, the business logic and requirements described in the interpretation make it critical for following process design.

2.2 Policy Representation

For deploying the business logic and requirements into the ICT systems of the organization, the interpretation is further modeled during knowledge engineering [11]. According to Boer et al. [12], knowledge models such as Web Ontology Language (OWL) models can be built based on sources of text, and traceability can be maintained between the models and the texts. Knowledge acquisition from text is a complicated and well-established knowledge engineering technique, and it is out of the scope of this paper. We focus on the challenge to connect ontology models, e.g., those represented by OWL, to BRs and SWSs in operational processes. Besides, the recent development of BRs in the Semantic Web domain focuses on a standard representation of BRs, e.g., the Rule Interchange Format (RIF) from W3C.

2.3 Administrative Processes

Automatic operational processes involve the execution of BRs that directly control process behaviors. Operational rules describe how an operation should be done. Operational BRs are subdivided into event-condition-action (ECA) rules and production rules. ECA rules are rules of the form "ON Event IF Condition DO Action". Production rules are rules of the form "IF Condition DO Action". Despite various rule technology solutions, rules are usually only applied in decision making in an automated process. Weigand et al. [13] have pointed out that many Business Rule approaches are restricted as they "concentrate on the application for decision-making and do not address the business process constraints that underlie service composition". In a SOA environment, this means the BR is hard coded in process, like BPEL. Although BPEL is the de facto standard for business process descriptions, BPEL is not flexible [13]. Service composition in BPEL relies on hardcoded BRs. This results in higher complexity of maintaining and updating processes and BRs, which is often a labor-intensive job. Furthermore, once the policy has been manually interpreted and translated to BRs and if it has been mixed with other sources of information, process compliance becomes ambiguous.

3 Guidelines for Creating Administrative Processes

Administrative organizations desire flexible business processes and must ensure that the business processes comply with policy. Concerning that many administrative organizations have BR and SOA infrastructure that support the second and third levels of the previous section we formulated the following guidelines for those organizations to use BR and SWS for dynamic creation of business processes. Those guidelines allow a clear coordination of the three levels, which is conceptualized in Fig.1. This integration will result in dynamic creation of business processes which can flexibly respond to the changes of policy and ensure the compliance all the time.

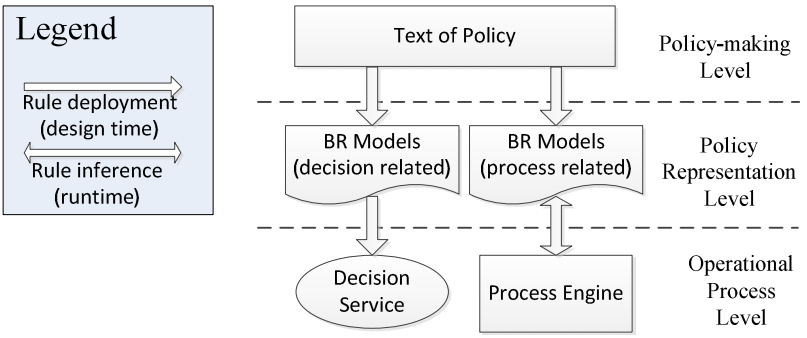


Fig. 1. Coordinating the components at the three levels

Guideline 1: Develop a Domain Ontology to Create Interoperability between SWSs and BRs.
Studies about the integration of Web Service and Business Rule technologies mainly focus on using BRs in the service composition. Orriëns et al. [14] have indicated that BRs can be used to determine how the composition should be structured and scheduled. BRs and Web Services were not compatible because they use different ontologies. Although they both can be semantic, they are not having a collective semantics foundation, as the same term/object can have different properties and a different relationship with other terms in different ontologies. Without a unified ontology, BRs for service composition had to know the exact interface of related services instead of their semantic description. This renders the semantic level of Web Services meaningless. D'Mello et al. [15] argued that rule driven composition requires BRs and SWSs to be documented in a unified ontology for successful composition. Besides, many BRs are originated from sources of law. The consistency of those BRs has to be concerned for compliance with policy.

Guideline 2: Distinguish between Decision and Supportive Services.
In a service-oriented system, services can be invoked from anywhere and managed independently. This allows the encapsulating of all forms of resources (provided by either humans or computers) into services and Semantic Web Service technology is used to describe them. For administrative organizations, we suggest defining two kinds of Business Services. The term "decision service" is used to describe a self-contained callable component with a view of all conditions and actions that need to be considered in making an operational business decision [16]. Decision services refer to policy, which implies that the use of these decision services is mandatory and they are the main services that can satisfy the service request with final decision making. Containing the policy related decision making logic, it is an essential step of the process to comply with legislation enforcement, but might not be able to satisfy the request on its own. In the execution of the decision services, other supportive services might need to be composed with it. In contrast, supportive services are not derived from policy but depend on other concerns of the organization. Supportive services can be used as infrastructural services with high reusability. Services like information retrieval services have a purpose to provide information/data for decision making. They allow a single information source to be reusable for multiple decision services. Other supportive services may facilitate decision services by providing user interfaces or report generation. They should be designed to be reusable as much as possible.

Guideline 3: Separate Decision Related and Process Related Business Rules.
BRs for decision making are widely used in administrative organizations. Yet, the contribution of BRs is not limited to decision making. In a SOA environment business rules is also able to control the composition of services into processes. Creating process related business rules avoid the logic of process flows being hardcoded in the applications (normally within a process unit for connecting the sequenced tasks). Distinguishing decision and process related BR allows a clear traceability from the policy making level to operational process level (see Fig. 1).

The separation of decision and process related BRs allows the decision services to encapsulate decision related BRs. A clear allocation of decision related BRs enable a fast response of changes in policy. Identifying process related BRs can avoid hard-coded BRs in process engine. This allows easier adaption of process flows.

The above three guidelines aim to facilitate the design of Business Process Management systems for administrative organizations. Those guidelines are technology independent and allow the organization to improve their current BR and SOA infrastructure. In the next section we will use an illustrative case study to demonstrate how to use those guidelines to create business processes for implementing policy.

4 An Illustrative Case Study

The Immigration and Naturalization Service (Immigratie en Naturalisatie Dienst, or IND) handles the admission of foreigners in the Netherlands. It is responsible for the execution of a complex set of regulations coming from different policies, including international law, national law, case decisions and so on. IND makes a large number of decisions, i.e. some 300.000 a year in the areas of asylum, standard objections to decisions, and naturalization. Under pressure from frequently changing policies, IND is one of the administrative organizations that have a desire for flexible and compliant business processes. In this chapter, we use the guidelines described in the previous chapter to show how they can help IND to realize this goal.

4.1 Adaptive Policy Scenario

A use case dealing with a Highly Skilled Migrant (HSM) policy that applies for a residence permit was selected to demonstrate the use of our guidelines. The policy is introduced to enable qualified foreigners to work in the Netherlands. In 2007, the annual income limitation on an applicant is at least 46.541 EUR or 34.130 EUR if the applicant is under 30. In 2008 there was a change in policy to encourage foreign intelligent graduates to work in the Netherlands. The foreign graduate who obtained a Bachelor or Master Degree at an accredited Dutch educational institution within one year before becoming employed, the income limitation was changed to 25.000 EUR annually.

In 2009 there was another change in legislation. Employees must have a gross annual income of at least 49.087 EUR, or 35.997 EUR if they are under the age of 30. There are two different situations to which the reduced wage criteria apply (25.800 EUR gross a year). The first situation is aimed at graduates that obtained a Bachelor or Master Degree at an accredited Dutch educational institution, similar to the 2008 situation. The second situation concerns Master and Ph.D. students who graduated in the Netherlands or at a university listed in the top 150 of two internationally recognized rankings.

4.2 Building a Domain Ontology

In our modeling approach, the first step is to define a domain ontology to provide consistent understanding of concepts. In Fig. 2 most concepts used in service description are involved and connected. For space reason, we didn't include every concept as this figure is enough to demonstrate the structure of a domain ontology.

As a typical case system, the domain ontology contains the identification of different clients under the *Person* class. We also defined an *Action* class, and its two subclasses with different semantic meanings. *Check* is for the provision functionality providing data by a supportive service. *Decide* indicates a decision service for providing a decision-making. *Resident Permit* and its subclasses sound like services. But in this example, they are final deliverables of processes instead of services. Therefore, they are defined as entities in the ontology model. Moreover, the data property of an entity (e.g. "person" has "ID" and "date of birth"; "applicant" has "income"; "graduate" has "institute" and "degree", "graduation date"; "decision" has "result") might eventually need to be included into the domain ontology for operational reasons. To improve the visual appeal, data properties are not presented in the diagram.

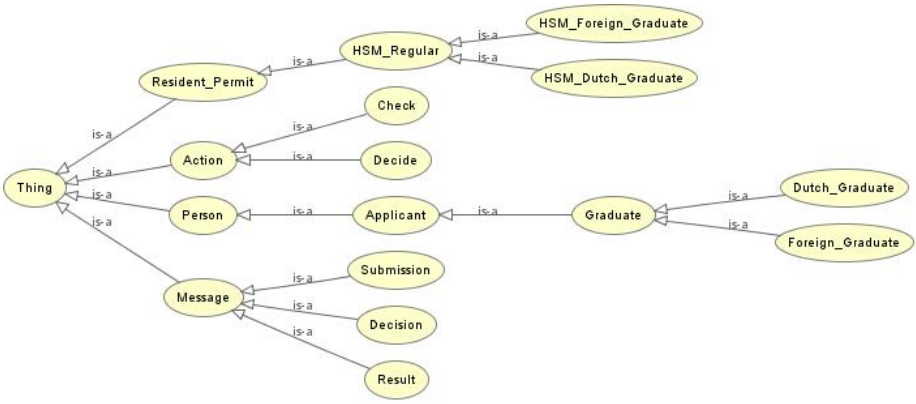


Fig. 2. Illustrative domain ontology in the HSM case (created using Protégé)

4.3 Defining Business Services

Using those concepts provided by the domain ontology model, we listed the designed services that used in HSM processes and their semantic representation in the following table. BSs are redefined using the criteria of Parnas [17], who suggested that each module should hide some design decision from the rest of the system. To simply express these services, we adopt a SWS description containing its functional semantics and the I/O information, so that a Web service can be specified as {service functionality, input set, output set} [18]. These services are summarized in the following table. Due to space reason, please refer to [8] for further detail of SWSs and domain ontology specification.

We assume that the system needs three decision services to decide whether an application can have a HSM resident permit. We need three different DS for this, because the policy distinguishes regular applicants, Dutch graduates and foreign graduate with different criteria for decision-making. More DS can be added if further exceptions are issued. Manual decision-making task can also be added as a new DS into the system to allow the control of special cases in the system. In this example, we only involve automatic DSs.

The system also needs several supportive services. To start the process, the system needs at least one supportive service (SS01) for submission intake. Alternative intake services can be issued to allow multiple submission channels (e.g. one for electronic channel, and one for front offices). We also need a supportive service (SS02) to check the income of applicants, a supportive service (SS03) to check the Dutch educational institution if the applicant is a Dutch graduate, and a supportive service (SS04) to check the top 150 ranking if the application is the applicant is a foreign graduate. Furthermore, we need one supportive service (SS05) for informing the application about the final result of decision making, namely whether he or she will receive a resident permit.

Table 2. The example of services in the HSM case

SWS	Specification	Description
Decision Services		
DS01	{{Decide, HSM}, { Income, Date_of_Birth}, {Decision}}	Decide whether the application satisfies regular HSM requirement
DS02	{{Decide, HSM_Dutch_Graduate}, { Income, Date of Birth, Degree, Result (Institute), Graduation_date}, {Decision}}	Decide whether the Dutch graduate satisfies HSM requirement
DS03	{{Decide, HSM_Foreign_Graduate}, { Income, Date_of _Birth, Degree, Result (Ranking), Graduation_date}, {Decision}}	Decide whether the foreign graduate satisfies HSM requirement
Supportive Services		
SS01	{{Check, Income}, {ID}, {Income}}	Check the income information
SS02	{{Check, Institute_Accreditment}, {Institute}, {Result (Institute)}}	Check whether an institute is accredited Dutch educational institution
SS03	{{Check, Ranking}, {Institute}, {Result (Ranking)}}	Check whether an institute is listed in the top 150 of a certain ranking

All the vocabulary used in SWS description and also the BRs descriptions in the next section should be involved in a domain ontology model, because the use of those vocabulary and terms must be consistent to allow understanding between components.

4.4 Defining Business Rules

To ensure compliance with policy, BRs that are used in the decision service are separated from the BRs that describe the process logic. Rules used in decision service are closely related to the legislation. In the HSM case, the legislation directly indicates the hard requirements on the age and income of applications. This (piece) of legislation should be connect to the rule to enable traceability. Being traceable to the piece of legal text allows easy maintenance of the rules, if any changes in legislation can impact them. Many expert systems can fulfill such a requirement. The BRs used by Web services are mainly productive rules. Their forms can be various depending on the implementation technology of the Web services. In our prototype, BRs are represented in RIF-Core [19]. An example of rules used in decision service DS01 is given as follows. This rule strictly responds on the HSM regular application according to the criteria of 2009. Note that most of the IRIs used in the examples of this paper are fictitious and do not represent real entities. To improve the visual appeal, the example takes advantage of the syntactic shortcuts (allowed by the RIF Datatypes and Built-ins specification). For instance, `100` is used instead of "`100"^^"xs:integer`".For the same reason, we directly use the mathematical and binary relations operations. For example, in practice, `?a>10` should be defined as `External(pred:numeric-greater-than(?a 10)`, where `pred` is the prefix defined as `Prefix(pred <http://example.com/rif-builtin-predicate#>)`.

```
Prefix(func <http:// example.com/rif-builtin-function#>)
Prefix(act <http://example.com/action#>)

Forall ?decision ?income ?date_Of_Birth(
  If  And(?decision # act:decision
          Or(?income >= 49087
              And(?income >= 35997
                   External(func:age(?date_Of_Birth)<30)
             )
         )
   Then Do(Modify(?decision[result->'granted']))
)
```

To create dynamic processes, we also need process related BRs, namely the BRs that are executed by the process engine. The crux is indicating which decision service should be invoke by process engine to deal the applications they are processing. As the functions of those decision services have been semantically described, the BRs then should indicate which decision service is for which kind of applications. We use the following RIF-core rules for giving an example. In this example, the Base directive provides yet another shortcut. The Base directive expands `<hsm_Regular>` IRIs to `"http://example.com/service#hsm_regular"^^rif:iri`.

```
Prefix(act <http://example.com/action#>)
Prefix(per <http://example.com/person#>)
Base(<http://example.com/service#>)

Forall ?Applicant(
  act:decide(?Applicant <hsm_Regular>):- ?Appli-
cant#per:applicant
  act:decide (?Applicant <hsm_Dutch_Graduate>):- ?Appli-
cant#per:dutch_graduate
  act:decide (?Applicant <hsm_Foreign_Graduate>) :- ?Ap-
plicant#per:foreign_graduate)
```

4.5 Testing Process Creation

In process creation, the following figure demonstrates the interactions among different system components. Process Engine first interacts with clients and receives HSM application submissions. To process the submission, the Process Engine invocates related Web services to make use of resources or knowledge for decision making. Finally, the decision result will be returned to the clients. The invocation of decision services is steered by the process related BRs. The selection of supportive services is based on the matching according to the semantic description of their functionalities and IOPEs.

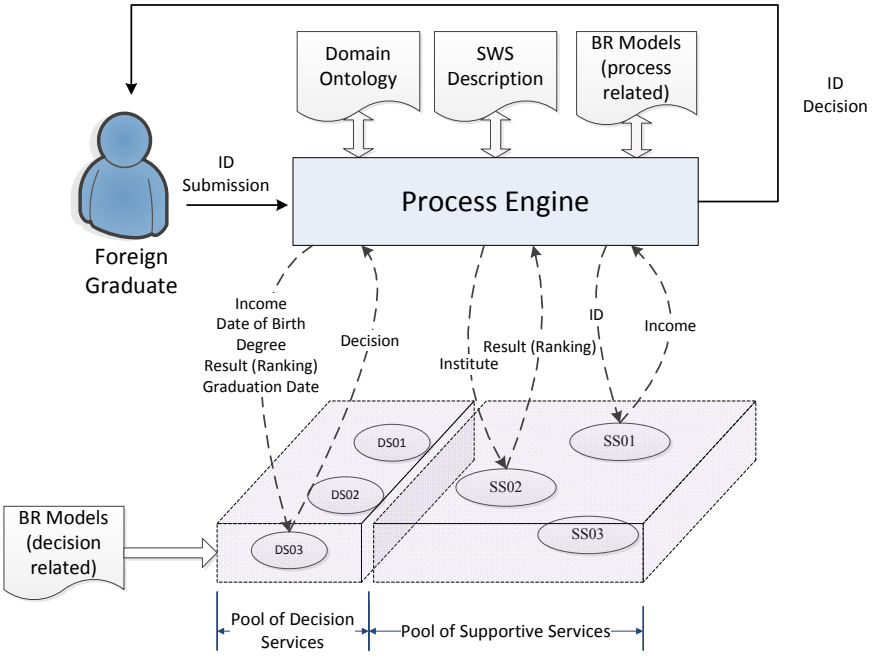


Fig. 3. An example of process creation for HSM foreign graduate applicant

To evaluate our guidelines, we invited IT architects, legal experts and process managers from IND to review the guidelines and the case study. Based on their experience, the IND experts evaluated our guidelines and case study by comparing them with their current way of business process management. The evaluation shows the added value of the guidelines including:

1. Faster adaptation of business processes to adopt changes in policy. As decision related BRs are allocated into decision services. It is easier to track the impact of changes in policy. The creation of processes is conducted by process related BRs. Adapting and testing those BRs is easier than reconfiguring the process engine.
2. Easier check for compliance with policy. As hardcoded and pre-defined processes can be avoid, checking the consistency between BRs and policy interpretation would be easier than checking the consistency between hardcode processes and policy interpretation. If any mistake happens, conflicting logic is easier to find out in BRs models than in hardcoded processes.
3. Clear integration of SWS technology. Being aware of how SWSs is used in process modeling facilitates the adoption of them.

5 Conclusion and Future Work

To satisfy requirements from frequently changed policy, administrative organizations desire adaptive and compliant policy implementation. This requires coordination of policy-making, policy representation and operational process level, and dynamic creation of operational processes. For having adaptability and compliance, the organization needs to coordinate functionality in process creation, and ensure communication between the elements. The proposed guidelines describe how to use BR and SWS technology to create administrative processes. A domain ontology enables communication between SWS and BR to achieve interoperability between those technologies. This allows them to share the same interpretation of the terms used in different components. Web services are used as the way to access and manage resources of the organization. SWS technology is used to describe the services such that BRs can understand them. Dynamic processes are created by process engine according to the BRs and using Web services as an ingredient. As BRs are not hardcoded in processes, the compliance to policy will be easier to check. Administrative organizations that implement frequently changed policy and would like to integrate multiple technologies to achieve higher flexibility in their business processes can benefit from those guidelines.

For future research, our guidelines provide a foundation for developing flexible and adaptive architectures and processes. The use in cross-organizational processes and supply chain management can be explored.

References

1. Weske, M.: Introduction. In: Kuropka, D., Staab, S., Tröger, P., Weske, M. (eds.) Semantic Service Provisioning. Springer, Heidelberg (2008)
2. Wang, H.H., Gibbins, N., Payne, T.R., Redavid, D.: A formal model of the Semantic Web Service Ontology (WSMO). Information Systems 37, 33–60 (2012)
3. Sabol, T., Furdík, K., Mach, M.: Employing Semantic Technologies for the Orchestration of Government Services. In: Vitvar, T., Peristeras, V., Tarabanis, K. (eds.) Semantic Technologies for E-Government, pp. 47–74. Springer, Heidelberg (2010)
4. Morgan, T.: Business rules and information systems: aligning IT with business goals. Addison-Wesley, Indianapolis (2002)
5. Lu, R., Sadiq, S.: A Survey of Comparative Business Process Modeling Approaches. In: Abramowicz, W. (ed.) BIS 2007. LNCS, vol. 4439, pp. 82–94. Springer, Heidelberg (2007)
6. Van der Aalst, W.M.P., Weske, M., Grünbauer, D.: Case handling: a new paradigm for business process support. Data & Knowledge Engineering 53, 129–162 (2005)
7. Lienhard, H., Künzi, U.-M.: Workflow and Business Rules: a Common Approach. In: Fischer, L. (ed.) Workflow Handbook 2005, pp. 129–140. Future Strategies Inc., Lighthouse Point (2005)
8. Gong, Y., Janssen, M.: Creating Dynamic Business Processes Using Semantic Web Services and Business Rules. In: Janowski, T., Davies, D. (eds.) 5th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2011). ACM Press, Tallinn (2011)
9. Gong, Y., Janssen, M.: From policy implementation to business process management: Principles for creating flexibility and agility. Government Information Quarterly 29, 61–71 (2011)
10. Hu, Y.-J., Yeh, C.-L., Laun, W.: Challenges for Rule Systems on the Web. In: Governatori, G., Hall, J., Paschke, A. (eds.) RuleML 2009. LNCS, vol. 5858, pp. 4–16. Springer, Heidelberg (2009)
11. Kordelaar, P., van Teeseling, F., Hoogland, E.: Acquiring and Modelling Legal Knowledge Using Patterns: An Application for the Dutch Immigration and Naturalisation Service. In: Cimiano, P., Pinto, H.S. (eds.) EKAW 2010. LNCS, vol. 6317, pp. 341–349. Springer, Heidelberg (2010)
12. Boer, A., van Engers, T., Winkels, R.: Traceability and Change in Legal Requirements Engineering. In: Casanovas, P., Pagallo, U., Sartor, G., Ajani, G. (eds.) AICOL Workshops 2009. LNCS, vol. 6237, pp. 74–92. Springer, Heidelberg (2010)
13. Weigand, H., Van den Heuvel, W.-J., Hiel, M.: Business policy compliance in service-oriented systems. Information Systems 36, 791–807 (2011)
14. Orriëns, B., Yang, J., Papazoglou, M.P.: A Framework for Business Rule Driven Service Composition. In: Benatallah, B., Shan, M.-C. (eds.) TES 2003. LNCS, vol. 2819, pp. 14–27. Springer, Heidelberg (2003)
15. D'Mello, D.A., Ananthanarayana, V.S., Salian, S.: A Review of Dynamic Web Service Composition Techniques. In: Meghanathan, N., Kaushik, B.K., Nagamalai, D. (eds.) CCSIT 2011, Part III. CCIS, vol. 133, pp. 85–97. Springer, Heidelberg (2011)
16. Taylor, J., Raden, N.: Smart (Enough) Systems: How to Deliver Competitive Advantage by Automating Hidden Decisions. Prentice Hall Press (2007)
17. Parnas, D.L.: On the criteria to be used in decomposing systems into modules. Communications of the ACM 15, 1053–1058 (1972)

18. Shin, D.-H., Lee, K.-H., Suda, T.: Automated generation of composite web services based on functional semantics. Web Semantics: Science, Services and Agents on the World Wide Web 7, 332–343 (2009)
19. W3C: RIF Core Dialect (W3C Proposed Recommendation 22 June 2010). World Wide Web Consortium (2010), http://www.w3.org/TR/2010/REC-rif-core-20100622/

Designing Visual Exemplars of 3D Online Store Layout Types

Ioannis Krasonikolakis[1], Adam Vrechopoulos[1], Nancy Pouloudi[1], and Katerina Goula[2]

[1] Athens University of Economics and Business, Athens, Greece
{krasos,avrehop,pouloudi}@aueb.gr
[2] National Technical University, Athens, Greece
katerinagoula@hotmail.com

Abstract. This paper presents the design issues in the visualization of five distinct store layout types in the context of 3D online retailing and discusses appropriate design decisions. The development of the stores is based on the requirements (layout characteristics) elicited from a three-round Delphi method with 3D expert users and designers which served as the qualitative empirical research vehicle. Along with the visualization of the characteristics of each layout type, the theoretical and practical implications, limitations, and the future research avenues of the study are discussed.

Keywords: Store Layout, Design, 3D Online Retail Stores.

1 Introduction

There has been considerable research in traditional and 2D online retail environments in light of the influence of store design/layout on consumer behaviour [5]. Levy and Weitz [8] illustrate in their textbook the three established layout types in traditional retail stores; i.e., the grid, the freeform, and the racetrack. Vrechopoulos et al. [15] transformed these types in the context of 2D online stores and found significant differences among them in terms of their influence on consumer behaviour. However, while retailing activity is gaining momentum in 3D online environments in recent years, there is not established knowledge about the layout/design formed in these stores.

Motivated by the aforementioned discussion, this study is based on the outcome of a qualitative research to visualize the design of the 3D online retail stores. A Delphi method was used to identify whether there are distinct layout types in 3D online stores and which are their characteristics, since this method is considered as an appropriate one for developing frameworks and concepts [10]. Following three rounds (13 participants in the first round, 10 in the second and 8 in the third one from both academic and business contexts), the Delphi method led to the identification of five distinct layout types and their underlying characteristics.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 311–324, 2013.

Elaborating on these results this study is positioned in the area of Electronic Retailing and follows a multidisciplinary approach to visualize the five distinct store types. Expertise from the research domains of Information Systems, Marketing, Architecture and Graphics Design was necessary and was combined to accomplish the aim of this research endeavor (i.e. research idea generation, design principles, implementation tools, etc.).

2 Development Options

Three development options were considered as the most appropriate to meet the requirements of the distinct 3D store layouts. The characteristics of each option, and the arguments for the most suited for the scope of this study are discussed below.

- Development of "real" 3D online retail stores

A possible development option could be the development of the five layout stores within a virtual world or in five different URL links on the internet. However, budget constraints indicated that this could not be an option. Specifically, in order to rent virtual land in a virtual world for five stores allocated in different virtual places or in five-storey block, one has to pay at least 500 US$ for a one month period. Also, specialists who have expertise with the specific development tools of the virtual world should be engaged in the development of the stores. Quite the same applies to the development of the stores in URL links.

- Visual exemplar of 3D online retail stores

Another development option could be the use of a 3D tool for the development of stores and a video recording could capture all aspects of the in-store layout patterns. There are freeware programs such as Goggle SketchUp providing the ability to develop a 3D appearance of a building. This option facilitates the development of 3D stores in a laboratory setting and provides a clear view of the interior of a store.

- "Paper-and-Pencil" visualization

This option could also meet the requirements of this research and is adopted in many research studies [12],[11],[16]. In this approach, the ground plan of each store layout could be designed in paper in order to visualize the store layout.

The second option is preferable to the third for deriving a clear picture/representation of store layout. Taking into consideration the relevant resources that made the first approach impractical, the second approach was adopted. Google SketchUp v.8 served as the main tool for building and modifying 3D models, as it is free software usually used by architecture to design 3D buildings. This tool offers the additional advantage of import and export capabilities to other design programs.

2.1 Retail Sector Selection

Retailing activity is taking place in various areas of the 3D shopping channel. There are several examples of apparel, furniture, real-estate, and tourism retail stores. The most popular outlets in coming years will be malls and retail stores [6]. 3D stores selling avatars' clothing is a common phenomenon in virtual worlds, while there are many companies such as Land's End, Kenneth Cole, and Levi's [4] that facilitate virtual try-on, through technology coming from Bodymetrics [7]. Following this market evolution, the apparel industry was considered a pioneering sector to serve as the experimental part of this study.

2.2 Product Selection and Allocation of Merchandize

Second Life is considered one of the leaders in virtual worlds [13]. To finalize the list of product categories that would be offered in the 3D online experimental stores of this study, it was considered appropriate to use the search engine for places in Second Life and visit the ten top listed apparel stores. These places are likely to be the most crowded and "famous" places. However, an obstacle that this study had to face is that the products offered in the stores visited in Virtual worlds, could not be copied and used for the scope of this study, because of copyright permissions. To overpass this obstacle, it was decided to use products offered in the Database of Google SketchUp 8. However, the variety of products offered by this program is limited. The use of Adobe Photoshop CS6 assisted to design clothes which are based on the products offered by Google SketchUp, but look quite different. The final database of the products developed was entirely used in almost all store layout types. The variety of products designed and used in the laboratory stores are: Dresses, Trousers, Ladies coats, Shirts, Shoes, Skirts, T-shirts, Ties, Tuxedos, Ladies' bags, Suitcases, Complete outfits (A common trend for retailers in 3D online retail stores is to sell complete formal or casual outfits). Indicative examples are presented in Figures 1,2 and 3:

Fig. 1. Indicative example of designed dresses, jeans, shirts, and skirts displayed in laboratory store layouts

Fig. 2. Indicative example of designed ladies coats, suitcases, and tuxedos displayed in laboratory store layouts

Fig. 3. Indicative example of designed complete avatars' outfits displayed in laboratory store layouts

As far as the allocation of products within each store is concerned, specifications coming from the Delphi method results, determined merchandise allocation guidelines in each store separately. Specifically, at the first-round of the Delphi, a panelist introduced the theme-based and similar-based allocation of products; these terms were validated from all panelists in the second-round Delphi. Theme-based display of products refers to a practice where products of the same category (e.g., shirts) are located in the same place in a part of the store, one next to the other. Similar-based display refers to the practice where complete outfits (e.g., shirt, jeans, belt, socks and shoes) are displayed, the one next to the other. The first practice serves the ease of searching among products of the same category, while the second serves matching an outfit's parts.

2.3 Real Models as Promoters of the Virtual Store's Collection

The same brand name was used in each store layout type for internal validity purposes of the forthcoming laboratory experiments. Clothing displayed in each store was named "Winick's Collection"; it comes from the first name of one of the researchers' avatar in the Virtual World Second Life since 2008, which is Winick Ceriano. The practice of displaying images of real world models to promote products, in walls of a 3D store, is common in Virtual Worlds retailers' stores. Copying images from models appeared on the Internet or 3D stores was not an option because of copyright permissions. To avoid this difficulty, one of the researchers captured images with an iPhone 4s mobile device, from six individuals who approved for their images to be displayed around each laboratory store layout. The final editing of the images was developed through Adobe Photoshop CS6. The images of the models are presented in figure 4.

Fig. 4. Indicative images of models displayed in the walls of the laboratory store layout types

2.4 Currency Selection

A notable amount of virtual worlds such as Whyville and Second life retain their own local currency within the world [9],[3],[2],[14]. In some virtual worlds, inhabitants can engage their virtual funds to get for example extra privileges, equipment etc; while in others the virtual currency can be exchanged for real dollars or Euros. For the virtual laboratory stores of this research study, it was decided that prices should be presented on the displayed products in order to provide a more realistic simulation of the shopping experience in the forthcoming laboratory experiment. In this regard, an imaginary currency was invented, namely the "$GRT". The "$" is a familiar notation of currency for participants, while the "GRT" symbolism comes from the consonants of the word "Gortyna" which is one of the most historic cities of the island Crete, achieving considerable growth three thousand years ago. A gortynian coin was auctioned at the end of 2011, for about 480.000 US$, which has been a world record for this type of auctions [1].

3 Graphical Interface of Store Layout Types

This section is dedicated to the presentation of images of the laboratory store layout types according to the specifications that resulted from the Delphi method. Specifically, there are three images representing views of each store layout type. It should be clarified that the above images on their current state do not include interactive features but rather transfer the layout design characteristics from text in to a graphical format. However, these prototypes enable the user to obtain a thorough view of each store through an automated navigation and display of its contents. Similarly, it should be also noted that it is out of the scope of the present paper to proceed in to the evaluation phase of the described layout types.

3.1 Store #1 (Medium Size Store)

The specifications of this layout type according to the Delphi study results are listed below, while the corresponding visualization is presented in Figures 5, 6 and 7:

- Theme-based/Similar-based display of products.
- Demo products or models wearing the products/images posted will help the customer reach a decision.
- Posters need to highlight the details of the products
- Insertion of screens in the floor plan to increase the amount of the display space they have.
- Requires avatars to move through the store rather than just being able to pan the walls with the camera.
- These stores tend to use images on the walls and may also use additional structures, but will leave some room in the middle for a model or two.

Fig. 5. View of a demo model

Fig. 6. Screens in the floor plan to increase space

Fig. 7. Images on the walls leaving space for model display

3.2 Store #2 (Large Warehouse Store)

The specifications of this layout type are listed below, while the corresponding visualization is presented in Figures 8, 9 and 10:

- Helpful display to the customer to compare products to each other.
- Functionality of comparing similar products.
- Theme-based/Similar-based display of products.
- Designer should be able to be contacted for further info on the products, because of the way they had the products designed.
- Ability to teleport into specific product-related areas.
- Easy ability to get into the building through alternative entry points.
- A virtual salesmen could guide customers find the products.

Fig. 8. Customer Service Kiosk (Contact designer info/compare products)

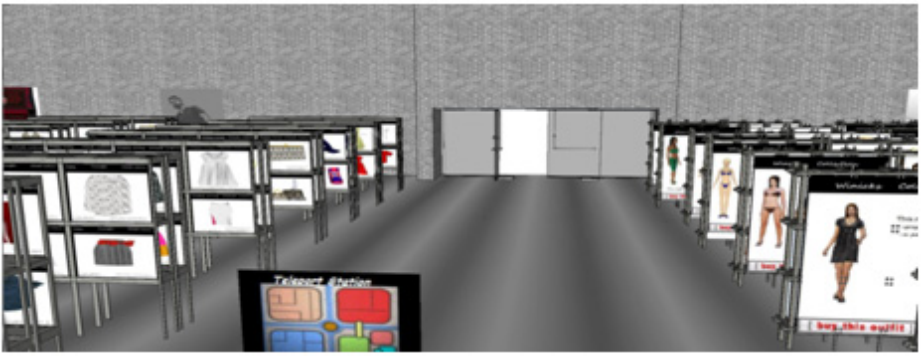


Fig. 9. Teleporting Station

Fig. 10. Salesman

3.3 Store #3 (Image-reliant Store)

The specifications of this layout type are listed below, while the corresponding visualization is presented in Figures 11, 12, and 13:

- Wall-only-items.
- Image stores are a great way for the retailer to reduce the lag of the store

- Theme-based display of products.
- Very simple product management for the end-user.
- Due to simple images, the simulation is much lighter and system requirements can be kept much lower. However, this sacrifices the realism of having a proper 3D model on screen.
- Inexpensive Approach: Makes it possible to show a lot of different items in what can be a relatively small space, especially when extra display walls are included.

Fig. 11. "Wall" items

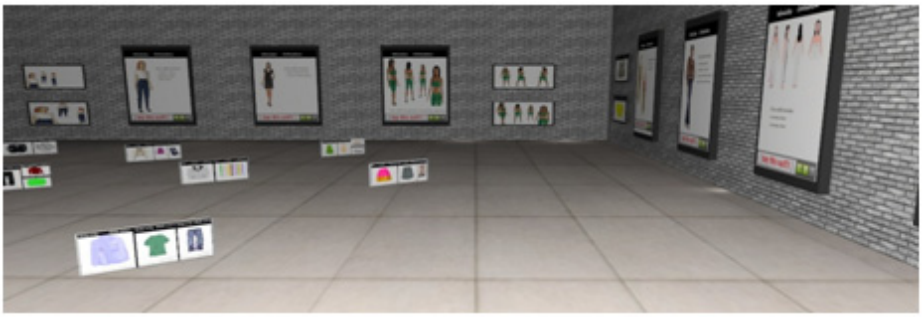


Fig. 12. Sacrifice of 3D models/simple images

Fig. 13. Inexpensive approach with extra display walls

3.4 Store #4 (Boutique Store)

The specifications of this layout type are listed below, while the corresponding visualization is presented in Figures 14, 15, and 16:

- They sell small items such as virtual hair for avatars or shoes.
- They tend to mimic physical stores with display cabinets and shelves.
- Customers browse the store quickly and if they don't find something they like, they can simply move on to the next one.
- The owner may also design note cards that are easy to give away and be shared between avatars/customers.
- Demo products also play a major role in this category.
- One should be able try on the product before reaching the decision of buying it.
- Clear display of products.
- Limited number of the available products.
- Feasible for some products such as artistic items.
- Theme-based/Similar-based display of products.
- Visual interest: interesting architecture, walls interesting architecture, wall of glass of glass, attractive materials – appeals to residents.
- Need to have enough blank space to make it easy for people to see the content of the shelves.
- Need to give distinctive names to items for people to be able to differentiate among them.

Fig. 14. Demo products

Fig. 15. Distinctive names to items

Fig. 16. Attractive materials/appeals to residents

3.5 Store #5 (Department Store)

The specifications of this layout type are listed below, while the corresponding visualization is presented in Figures 17, 18, and 19:

- Ability to find a great variety of products in a specific place (e.g. from clothing to food).
- Similarities to traditional stores regarding space layout.
- Similarities to traditional stores regarding product clustering.
- Similarities to traditional stores regarding store's walk-through scenarios.
- Simulation of traditional (physical) Department retail stores.

Fig. 17. Similarities to traditional stores regarding space layout

Fig. 18. Simulation of traditional allocation of products

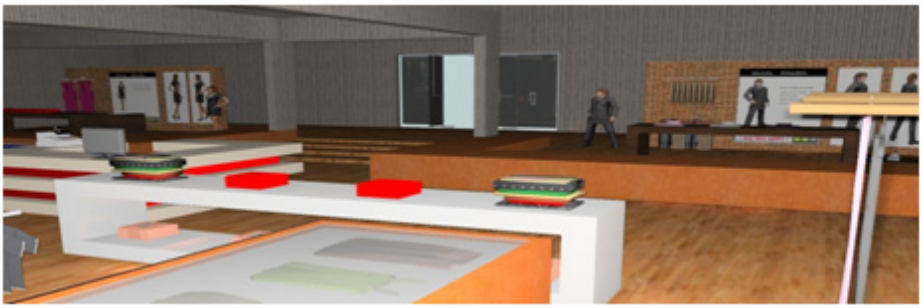


Fig. 19. Similarities to traditional stores regarding product clustering

4 Implications, Limitations and Avenues for Future Research

The main theoretical contribution of this paper refers to the visualization of the 3D retail stores' layout types similarly to the established existing theoretical knowledge in conventional and 2D Internet retailing. Along these lines, the exemplar of the five distinct 3D online store layouts serves the visual representation of the stores and can be critical tool for marketing research initiatives (e.g. investigate store layout effects on consumer behaviour). While the visualization and underlying store layout characteristics have been adequately illustrated in the traditional and 2D online retail context, this is the first study to address and develop these issues in the context of 3D online environments through a multidisciplinary research approach. However, it should be reminded that the research objective of this paper was not to present the methodology followed for developing the layout classification scheme but rather to present the methodology followed for visualizing the corresponding layouts' design patterns.

A limitation of the study is that the store types were not developed within a virtual world that would be useful for the design and execution of field experiments that ensure higher external validity compared to laboratory ones. However, apart from the budget constraints that averted this venture, the followed approach has an additional advantage in that it is not subject to potential brand name effects and other field experiment's related phenomena that could bias the results of a causal research study. In other words, the followed approach ensures higher internal validity of the forthcoming experimental testing. Another limitation of the study is that the layout types were derived through business and academic experts' feedback and not from users-customers. Thus, future research could proceed to testing this framework against real customers.

This visualization could be used for educational purposes, assisting academia and practitioners to get a picture of how the current business practice, as far as store design is concerned, is formed in 3D online environments. Similarly, the methodological steps taken for visualizing the store layouts could also serve as a useful instrument

for replicate studies. As also discussed above, this exemplar could be used as a research tool in order to conduct experiments in a laboratory setting, to identify how each layout store type influences consumer behavior. Specifically, this visual representation can help research the link between layout, customer experience, e-word of mouth, and online purchase intentions. Finally, a more technical future research approach could enrich the visualization schemes of the present study with more interactive features and dynamic capabilities.

References

1. Archeology Newsroom.: Ancient Cretan Coin Sold for 479.000 US$ (2011), http://www.archaiologia.gr/blog/2011/11/28/%CE%B1%CF%81%CF%87%CE%B1%CE%AF%CE%BF-%CE%BA%CF%81%CE%B7%CF%84%CE%B9%CE%BA%CF%8C-%CE%BD%CF%8C%CE%BC%CE%B9%CF%83%CE%BC%CE%B1-%CF%80%CE%BF%CF%85%CE%BB%CE%AE%CE%B8%CE%B7%CE%BA%CE%B5-%CE%B3%CE%B9%CE%B1-479/ (accessed at: November 2012)
2. Bainbridge, W.: The Scientific Research Potential of Virtual Worlds. Science 317(5837), 472–476 (2007)
3. Balkin, M.J.: Law and Liberty in Virtual Worlds. 49 N.Y.L. Sch. L. Rev. 63 (2004)
4. Faust, M.-E., Carrier, S.: 3D Body Scanning's Contribution to the Use of Apparel as an Identity Construction Tool. In: Duffy, V.G. (ed.) Digital Human Modeling, HCII 2009. LNCS, vol. 5620, pp. 19–28. Springer, Heidelberg (2009)
5. Griffith, D.: An examination of the influences of store layout in online retailing. Journal of Business Research 58(10), 1391–1396 (2005)
6. Krasonikolakis, I., Vrechopoulos, A., Christoforou, D.: An Empirical Investigation of Users' Characteristics in Virtual Reality Retailing. In: 2nd International Conference on Services Marketing, Thessaloniki (November 2009)
7. Laird, S.: Clothes Shopping With Bodymetrics Lets You Try It On for Virtual Size. Mashable Tech (2012), http://mashable.com/2012/01/09/bodymetrics-augmented-reality-shopping/
8. Levy, M., Weitz, B.A.: Retailing Management, 4th edn. McGraw-Hill, New York (2001)
9. Messinger, P., Ge, X., Stroulia, E., Lyons, K., Smirnov, K., Bone, M.: On the Relationship between My Avatar and Myself. Journal of Virtual Worlds Research 1(2), 1–17 (2008)
10. Okoli, C., Pawlowski, S.: The Delphi Method as a Research Tool: An Example, Design Considerations and Applications. Information & Management 42, 15–29 (2004)
11. Pettit, F.A.: A Comparison of World-Wide Web and Paper-and-Pencil Personality Questionnaires. Behavior Research Methods 34(1), 50–54 (2002)
12. Russell, M., Haney, W.: Testing Writing on Computers: An Experiment Comparing Student Performance on Tests Conducted Via Computer and Via Paper-and-Pencil. Educational Policy Analysis Archives 5(1) (1997)
13. Shin, H.D.: Understanding Purchasing Behaviors in a Virtual Economy: Consumer Behavior Involving Virtual Currency in Web 2.0 Communities. Interacting with Computers 20, 433–446 (2008)

14. Stuart, B.: Virtual Worlds as a Medium for Advertising. ACM SIGMIS Database 38(4) (2007)
15. Vrechopoulos, A., O'Keefe, R., Doukidis, G., Siomkos, G.: Virtual StoreLayout: An Experimental Comparison in the Context of Grocery Retail. Journal of Retailing 80(1), 13–22 (2004)
16. Wright, D., Aquilino, W., Supple, A.: A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use. Public Opinion Quarterly 62, 331–353 (1998)

A Multicriteria Methodology for the Selection and Prioritisation of Public Services

Anna Georgiadou, Panagiotis Kokkinakos, Dimitrios Panopoulos, Sotirios Koussouris, and Dimitrios Askounis

School of Electrical & Computer Engineering, National Technical University of Athens
georgiadou.anna@gmail.com,
{pkokkinakos,dpano,skous,askous}@epu.ntua.gr

Abstract. The pressing need to deploy the information and communications technology in the public sector effectively and efficiently along with the complexity of the decision making process have thrust research towards the development of public services prioritization models. The purpose of this paper is to outline a general multicriteria decision support method in order to identify the services that, if transformed to electronic ones, would act as ambassadors of the new era. Various technological, social, political and economic criteria have been taken under consideration. The proposed model is based on the Analytic Network Process (ANP), which takes into account both tangible and intangible criteria with interdependencies and feedback among them. A real case scenario concerning a Public Services selection process in Greece has been used to validate the presented model.

Keywords: Public Services, Analytic Network Process (ANP), Analytic Hierarchy Process (AHP), Multicriteria Decision Analysis (MCDA).

1 Introduction

The rapid technological development of the last decades has radically transformed the way government information and services are delivered to the citizens. Web-based telecommunication technologies have been utilized in order to improve and/or enhance the efficiency and effectiveness of service delivery in the public sector [1-2]. The general term "e-Government" was introduced so as to describe this technological penetration and public transformation and, today, it has progressed to the point where it has proven to be a force for effective governance and citizen participation, both at international and local levels [3].

This upward trend in e-Government development [4] has accelerated despite, or maybe in part also due to, the current financial crisis and the pressing need for governments to be competent, transparent, accessible and efficient. The compelling factor and the bitter early lesson learned from the crisis is that trust matters, and lack of confidence in government, as in markets, can lead to calamity. The capacity to respond under difficult conditions and deliver the expected results are cornerstones of effective government and a foundation of public trust.

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 325–337, 2013.

In this light, due to the aforementioned global economic crisis [5], the probability of a wrong investment has to be minimized and the spending of public money needs to be targeted and with low risk. In other words, the spawned need of public services prioritization will soon become a pressing challenge which could decide the future of e-Government's embracement and effectualness. The purpose of the present paper is to propose the implementation of a general multicriteria decision making method in the prioritization of public services. The question that is which of the public service scenarios under consideration should be delivered first, partially or fully, in an electronic way [6]. This approach will enable governments take into account, not only historical data, but also the current conditions, the trends and the tacit knowledge of the experts.

In general, the domain of Multiple Criteria Decision Analysis (MCDA) concerns the approach of explicitly taking into consideration the pros and cons of a plurality of points of view, in order to make a decision [7-9]. MCDA is an activity which helps making decisions mainly in terms of choosing, ranking or sorting the potentials [10]. The proposed model is based on the Analytic Network Process (ANP), an advanced MCDA method [11], and is meant to prioritize public services by using both quantitative and qualitative variables.

This paper's object is twofold. It emphasizes on outlining the criteria which need to be considered in public services selection, while, on the other hand, it offers a comprehensive approach on how ANP can be utilized to aid the decision making process in the aforementioned problem. Its practical value is illustrated via an example in the area of the current Greek reality.

The rest of this paper is organised as follows. The second section offers a brief literature review and corresponding analysis of the AHP and ANP methodologies and describes the service scenarios under consideration, while it also defines the criteria which were used for the multicriteria model introduced. Moreover, in this section, the entire methodology application is presented as well as the respective results. The paper concludes with the last section where basic findings are discussed, minor limitations of the approach proposed are underlined and possible future steps are recognized.

2 Proposed Approach and Application

The current financial crisis enforced in a way the commanding penetration of the web-based technologies into the public sector. In this rather restrictive financial reality, the sectors where e-Government would most benefit the government as well as the users of public services (meaning both citizens and businesses) should carefully be examined and scaled. Therefore, the authors have tried to summarize and pinpoint the most important criteria in evaluating the delivered services while proposing the use of a multicriteria decision making approach in this field.

2.1 Methodological Multicriteria Background for Ranking and Prioritising Services

The methodology employed for selecting and prioritising services is based on the Analytic Network Process (ANP), a multiple-criteria decision analysis method which, since its introduction by Thomas L. Saaty, has been used around the world in a wide variety of decision situations, in fields such as government, business, industry, healthcare, and education [12]. The ANP is a generalization of the Analytic Hierarchy Process [13-14]. The Analytic Hierarchy Process (AHP) is designed to structure a decision process in a scenario affected by multiple independent factors. Based on this methodology, a complex problem is divided into sub-problems organized according to hierarchical levels [15], with each level denoting a set of criteria or attributes related to the specific case. The top level of the hierarchy denotes the main goal, the bottom level contains the alternative or actions considered when achieving the goal, and the intermediate levels denote the factors that influence the upper levels. In this context, the AHP methodology makes it possible to compare different factors, where each factor's importance influences the final solution and the rank of the alternative that is to be used for reaching this solution [16-17].

2.2 Service Scenarios' Description

The proposed model applies the ANP to the problem of evaluating, prioritizing and selecting public services [18]. The model has been validated through a real case implementation [19] concerning the prioritization of five different public services delivered by the Greek public sector. These services have been selected with the contribution of the Ministry of Administrative Reform and e-Government which provided the public service scenarios which are under consideration for improvement and radical transformation of the way they are delivered. The five most important ones which are presented below:

1. *Access extracts of insurance record of Social Security Institutes*: According to almost all European social security systems, every employee (citizen) must be registered to a public security organization. For each working day, a part of employee's salary is paid to the social security institute so that the employee can receive social security services (e.g. hospitalization, allowances). Usually the employer (e.g. a company) keeps a part of employee's salary and adds the employer's contribution so as to pay a specific amount of money to the Social Security institute in charge. The employee can verify that the employer has paid its contribution to the Social Security institute by accessing and getting an extract of his/her record regarding the working days that have been paid for him/her. Currently this service is provided after an electronic or conventional application depending on the each specific case's sophistication level.
2. *Certificates of various civil acts*: Citizens request from e.g. a Citizen Service Centre the indented certificates while providing the necessary supporting documents (e.g. photocopy of the identity card for a birth certificate). This procedure has a wide application since such certificates are prerequisite to many other services (e.g.

changing municipality). The authorized employee of the Citizen Service Centre collects and checks the supporting documents. In case something is missing, the procedure halts and the citizen is informed about the derived issue, otherwise, the Citizen Service Centre employee completes the electronic application and creates a new folder in the Citizen Service Centre's informational system. Then, the employee sends the application along with the supporting documents to the proper authority (via fax, post office etc.). When the Citizen Service Centre receives the certificate from the authority, the citizen is notified to collect it from the desired Citizen Service Centre.

3. *Monitoring compensation procedures of Hellenic Organisation of Agricultural Insurances:* The Hellenic Organisation of Agricultural Insurances - ELGA is a governmental body for supporting agricultural production. ELGA provides social insurance to farmers and agricultural workers while, on the other hand, provides insurance on the agricultural production and compensates the producers in cases of physical disasters destroying their harvest. Whenever the harvest of a farmer is destroyed – on the condition that he/she had applied all the necessary measures to be able to grow his/her products with safe – he/she may claim for compensation. The proposed scenario has to do exactly with the procedure that is being followed by a farmer in order to get compensated after such a disaster. The service has as a main purpose to provide real-time information to the farmers concerning the stage of their requests for compensation, the height of the compensation decided and the time of the deposit. In addition the service provides to the farmer all the relevant certificates needed for his/her tax declaration concerning the compensation. The service is currently operated by the administrative services of ELGA with the involvement of the citizen service centers.
4. *Transferring citizen rights to another municipality:* The proposed scenario refers to the case that a citizen has moved to another municipality and wants to transfer his/her rights in order to be able to vote or to participate to several activities in the municipality he/she has moved.
5. *Monitoring an application submitted through e-Gov procedures:* Citizen Service Centers in Greece act as a mediator between the citizen and the competent authorities. Whenever a Citizen requests a service, Citizen Service Centres create a dossier of the case including the application forms and the supporting documents. This dossier is transferred to the competent authority. Citizen Service Centres are not responsible/authorized for the execution of the service. Usually an estimated time for the execution of the service is announced to the citizen. As long as the response of the competent authority is in the expected timeframe, the result of the case is transferred to the Citizen Service Centre. As a final step Citizen Service Centres contact the citizen announcing that his/her case has been completed and deliver the result to the citizen. A new service has to be designed in order so that the Citizen Service Centres and the Citizen have constant information regarding the status of their case. This service will allow immediate response to the Citizens for the status of their case. Moreover, this service will provide valuable information for the public administration regarding the steps of a process or problems that usually cause delays. This is a valuable tool for indenting areas for public administration reform and simplification of the procedures.

2.3 Criteria Definition

In our effort to develop a public service evaluation ANP model, the most critical point was to define the criteria which need to be taken into consideration [20]. Towards that direction, the authors performed an extended research work concerning evaluation criteria in the public sector and proceeded to selecting the common criteria that were common among the most relevant approaches [29-32] and reflect in the best way the cases examined. These criteria, combined with the ones extracted from the Services Description Template, introduced in [21], aim to document the various facts and details of every public service scenario analysed. The criteria were reviewed and categorized into clusters that were later on communicated and evaluated by the executive officers of the Greek Ministry of Administrative Reform and e-Government, hereafter referred as decision makers. This procedure concluded in the formation of the final model.

In the following table, a detailed description of the criteria used is presented combined with their categorization into clusters:

Table 1. Criteria Description

Cluster	S/N	Criterion	Possible Values	Description
SERVICE FLOW	A	Public (Final) Service	Yes / No	Indicates whether the service is public (final) or if it is an intermediate process.
	B	Self-appointed call of Service	Yes / No	Defines whether the service could be called only with the acquiescence of service receiver or, alternatively, automatically by a public organization.
	C	Re-usability by other services	High/Medium/Low	Indicates whether this service, although final, is also prerequisite to other public services – so it can be reasonably considered as more important
	D	Service Execution Domain	Internal/ External	Indicates whether this service is executed within the boundaries of an organization (e.g. a municipality), or it includes interaction with other, «external» organizations (e.g. the police) as well.
SERVICE COMPLEXITY	E	Legal Framework Implication	High/Medium/Low	States in which degree legislation is affecting the operation of a service.
	F	Level of compliance with National Interoperability Frameworks	Yes / No / No Interoperability Framework existing	Indicates whether the delivery of the service is compliant with the correspondent interoperability framework.
	G	Requirement for Personal Presence	Yes / No	Indicates whether physical presence at the submission of the application, the delivery or any other step of the service is required.

Table 1. (*continued*)

	H	Level of Support by IT systems	Full / Semi / No support	Indicates if the service is fully automated, if some steps are done manually, or if there is no support at all – depending on the maximum level of support the specific service could provide.
SERVICE IMPORTANCE	I	Based on European Policies	Yes / No	Indicates whether the service is aligned with the i2010 list of 20 Basic Public Services and the overall European eGovernment Action Plan. The list consists of 12 services for citizens and 8 for businesses.
	J	Existence of a pan-European dimension	Yes / No	Indicates whether a pan-European dimension of the specific service exists. An indicative list of public services which have a pan-European dimension and shall be taken into account in the present task includes, i.e.: Request and delivery of birth, marriage, death certificate, Services related to nomadic work, Declaration to the police
	K	Potential to deliver value within a certain short time frame	High/Medium/Low	It is vital to select a service that can deliver value within a specific short time frame. A selection of an interesting but time-consuming public service could lead to a failure of keeping the time frame.
	L	Importance for the Service Provider Organization	High/Medium/Low	States if the service is important and probably a part of the strategic plan of a public organization.
	M	Political dimension	High/Medium/Low	Indicates the compliance of the service to the Organisation's Mission and the Strategic Policies as well as the projected Public Perception of the service and its political impact [22].
SERVICE LOAD DATA	N	Total Service Duration	Duration in hours/days	Indicates how many hours/days does it take for the service to complete – no matter if the service is provided conventionally or electronically.
	O	Frequency	Requests/ year	Indicates the number of the requests of the specific service in a given amount of time.
	P	No of Inputs	Number	Defines the number of prerequisites of the service.

At this point, a reference should be made to the need to adjust the criteria proposed via the Services Description Template so as to generalize their application. More specifically, the criteria "Requirement for Personal Presence at the submission of the

application" and "Requirement for Personal Presence at the delivery of the service" were unified to lead to the more generic one "Requirement for Personal Presence" mainly due to the different nature of each public service scenario under investigation. Moreover, the Services Description Template seems to totally overlook the political dimension of a public service [23] downgrading its role to the decision making procedure in the public sector.

From a technical point of view, the structure of the ANP model is described by its clusters and elements and the connection between them. These connections indicate the flow of influence among the elements. The resulting model is illustrated in the following figure and consists of five clusters:

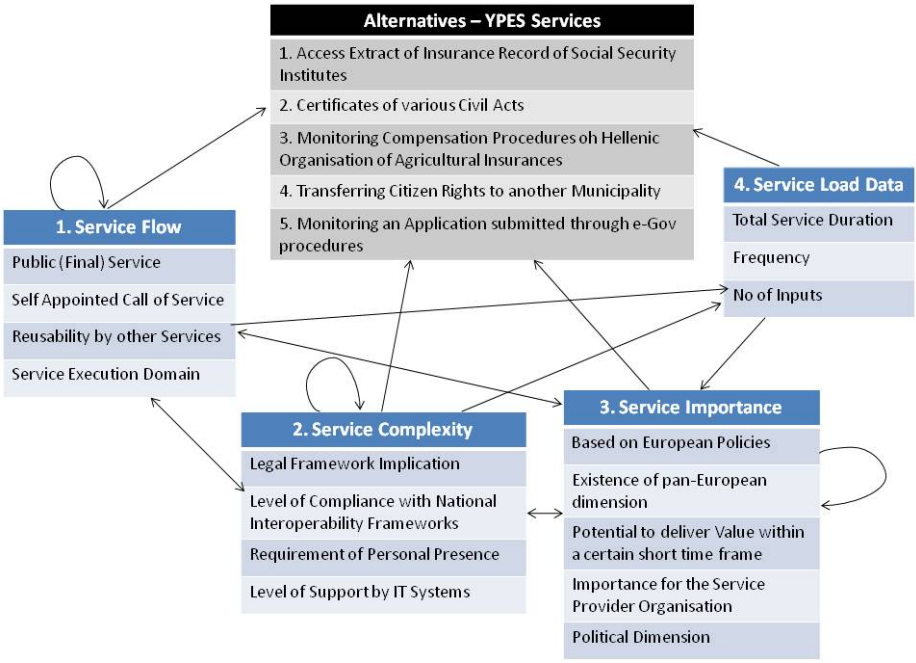


Fig. 1. The Public Service Evaluation ANP Model

For the ANP implementation, Super Decisions[1] software was employed, developed by the ANP Team (working for the Creative Decisions Foundation), to alleviate the mathematical burden. The figure depicted above represents the Public Service Evaluation ANP Model as constructed with the usage of Super Decisions software [24] so as to implement our example case study for the Greek public sector.

[1] http://www.superdecisions.com/

2.4 Analysis Approach and Paired Comparisons

Having decided to adopt the influencing analysis approach (recommended and broadly used in bibliography) and with the valuable contribution of the decision makers, the connections among the clusters and elements were defined (depicted in Figure 1 with arrows) to reflect the interaction of the different factors in the real environment. For instance, the bidirectional arrow from cluster "1. Service Flow" to cluster "3. Service Importance" captures the notion of outer dependence on ANP and is interpreted in means of influence among their elements. On the other hand, the concept of inner dependence is depicted with a feedback loop such as the one present in cluster "2. Service Complexity" and represents the influence among different elements of the same cluster.

Following the final formation of the public service MCDA model, the collaborating officers of the Ministry of Administrative Reform and e-Government were asked to respond to a series of pair wise comparisons with respect to a control criterion. "Super Decisions" software is able to calculate the inconsistency ratio for each comparison matrix, so the most consistent value for the entries can be determined. The inconsistency measure is useful for identifying possible errors in judgments as well as actual inconsistencies in the judgments themselves. Afterwards, the Clusters Priority Matrix is constructed; whose columns consist of the weighted priorities derived during the pair wise comparisons.

Table 2. Clusters Priority Matrix

	1. Service Flow	**2. Service Complexity**	**3. Service Importance**	**4. Service Load Data**	**Alternatives**
1. Service Flow	0.002652	0.558299	0.487794	0000	0000
2. Service Complexity	0.681786	0.121505	0.121957	0000	0000
3. Service Importance	0.079742	0.110766	0.195125	0.500000	0000
4. Service Load Data	0.233168	0.087925	0000	0000	0000
Alternatives	0.002652	0.121505	0.195125	0.500000	0000

The next step of the method was the computation of the SuperMartix, the Weighed SuperMatrix and the Limiting Matrix. These matrixes represent the impact or influence of each elements on other elements of the problem and are computed by the Super Decision software. Indicatively, the Weighted SuperMatrix of the examined problem is presented in the following table:

Table 3. The Weighted SuperMatrix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	1	2	3	4	5
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	1.000	0	0	0	0.333	1.000	1.000	0	0	0.500	0	0	0	0	0	0	0	0	0	0	0
C	0	1.000	0	0	0.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0.333	0	0	0	0	0.500	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0.500	0	0	0	0	0	0	0	0	0	0	0	0
F	0	1.000	0	1.000	0	0	1.000	0	0.500	0.750	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0.250	0	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	1.000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0.333	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0.500	0	1.000	0	0	0	0	0.333	0	0	1.000	0.500	0.167	0	0	0	0	0	0	0
L	1.000	0.500	0	0	0	0	0	1.000	0.333	0	0	0	0.500	0.833	0	0.500	0	0	0	0	0
M	0	0	0	0	0	0	1.000	0	0	0	0	0	0	0	0	0.500	0	0	0	0	0
N	0	1.000	0	0	0.333	0	0.500	1.000	0	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	1.000	0.333	0	0.500	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0.429	0.140	0.238	0.200	0.243	0.243	0.292	0.080	0.048	0.048	0.286	0.113	0.091	0.478	0.253	0.303	0	0	0	0	0
2	0.048	0.554	0.238	0.200	0.243	0.027	0.089	0.152	0.429	0.429	0.286	0.339	0.091	0.203	0.139	0.303	0	0	0	0	0
3	0.048	0.054	0.238	0.200	0.243	0.243	0.292	0.152	0.048	0.048	0.071	0.401	0.091	0.042	0.139	0.303	0	0	0	0	0
4	0.429	0.188	0.238	0.200	0.243	0.243	0.036	0.096	0.429	0.429	0.286	0.059	0.091	0.053	0.052	0.031	0	0	0	0	0
5	0.048	0.065	0.048	0.200	0.027	0.243	0.292	0.519	0.048	0.048	0.071	0.088	0.636	0.224	0.418	0.060	0	0	0	0	0

2.5 Final Prioritization Results

The final results, prioritization (normalized values) of public services under evaluation as derived from the limit matrix are presented in Table 4 and, for a more representative overlook of the outcome, in Figure 2.

Table 4. Final Public Services Prioritization Results

Alternatives	Total	Normal	Ideal	Ranking
1. Access, extracts of insurance record in Social Security Organisation	0.0700	0.2867	1.0000	1
2. Certificates of various civil acts	0.0532	0.2177	0.7593	2
3. Monitoring compensation procedures of Hellenic Organisation of Agricultural Insurances	0.0402	0.1647	0.5746	4
4. Transferring citizen rights to another municipality	0.0446	0.1828	0.6375	3
5. Monitoring an application submitted through e-Gov procedures	0.0362	0.1482	0.5168	5

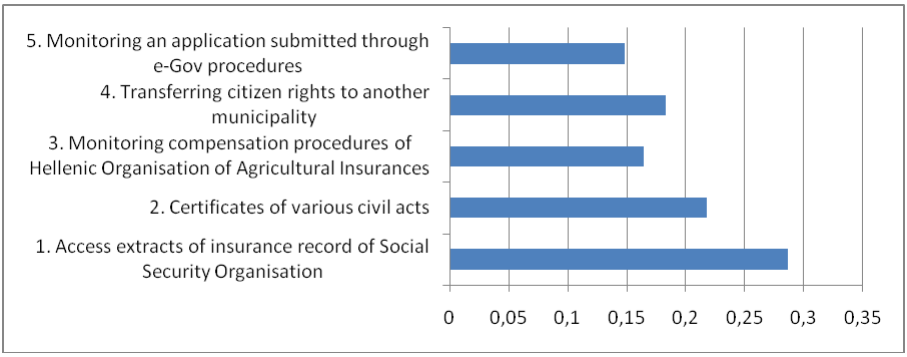


Fig. 2. Prioritization of Proposed Public Services

As shown above, the priorities of the services which were derived after applying the multi-criteria analysis do not represent only their importance for the Ministry of Administrative Reform and e-Government but also how much appropriate they are based on the technical implementation requirements. It becomes obvious that services like the prevailing one (Access extracts of insurance record of Social Security Organisation), which affect a large number of end users but are also compatible with the public authorities' initiatives and frameworks, take precedence against other scenarios that do not meet these standards.

Nevertheless, it needs to be noted that, in order to reach a justified and viable result, a time consuming and possibly tiring effort is needed both by the Public Sector (which has to provide all respective data) and by the supporting technical user (who operates the suitable software) [25].

2.6 Sensitivity Analysis

In order to check the stability of our results, a sensitivity analysis is performed with a series of a "what if" kind of questions to verify that the final classification is stable

when the inputs, whether judgments or priorities, are changed. Therefore, the authors have proceeded by radically altering the responses received during the pairwise comparisons.

The first and most important part of our sensitivity analysis was the one regarding the "Level of compliance with National Interoperability Frameworks", which is the dominant criterion based on our ANP analysis. The procedure showed that irrespective of its value, the rank of the public services scenarios is preserved.

The rest of our analysis consisted of similar sensitivity tests for the rest of the criteria used in our model and the results led to the conclusion that the outcome is sufficiently stable and does not change the overall rank [26].

3 Conclusions

The prioritization of public services is a rather difficult endeavour as it considers many conflicting criteria, both tangible and intangible. The present paper introduced an MCDA approach, based on the ANP methodology, capable of capturing this complex decision environment. In more detail, the proposed model enables the decision maker to better comprehend the various factors that influence the final outcome while, at the same time, documents the evaluation results in an indisputable way. The case study used to evaluate the effectiveness of the proposed model referred to a limited number of scenarios; however, the presented methodology is also applicable to sets of numerous alternatives with the cost of complexity. In our example, the highest priority was assigned to the "Access extracts of insurance record of Social Security Organisation" while the most dominant criteria were proven to be the "Level of compliance with National Interoperability Frameworks" and the "Re-usability by other services" demonstrating the importance service complexity and flow factors bear in this field.

The authors feel the obligation to underline the fact that the ANP methodology proposed bears a few limitations which mainly derive from the subjectivity of the input by the decision makers and the time-consuming nature of the pairwise comparisons which may lead to inconsistencies and, consequently, to doubtful and wrong results. Yet these limitations are not considered to be restrictive mainly due to the complexity of the e-Government field.

The present paper raises several important issues that could spark further research concerning the evaluation of the model and its implementation at additional case studies in different European countries.

In the near future, it would be useful if a coordinated pan European initiative among Public Institutions took place so as to define common and mutually accepted criteria for the prioritization of public services in order to formulate a unified model for the sophistication of public service delivery.

Last but not least, an ICT tool [27-28] could be developed in order to apply the aforementioned methodology efficiently and effectively to any kind of service providing useful results and guidelines to its final users.

Acknowledgements. Part of the work presented in this paper is sponsored by the European Commission in the frames of the FP7 project "COCKPIT - Citizens Collaboration and Co-Creation in Public Service Delivery" (Contract No: FP7-ICT-2009-4 COCKPIT, http://www.cockpit-project.eu/).

References

1. Jeong, C.H.: Fundamental of Development Administration. Scholar Press, Selangor (2007) ISBN 978-967-504-5080
2. Beynon-Davies, P., Martin, S.: Electronic Local Government and the Modernisation Agenda: Progress and Prospects for Public Service Improvement. Local Government Studies 30(2), 214–229 (2010)
3. United Nations: Department of Economic and Social Affairs, United Nations E-Government Survey (2010)
4. Millard, J.: Reorganisation of Government Back-Offices for Better Electronic Public Services. In: Traunmüller, R. (ed.) EGOV 2004. LNCS, vol. 3183, pp. 363–370. Springer, Heidelberg (2004)
5. Niemira, M., Saaty, T.L.: An analytic network process model for financial-crisis forecasting. International Journal of Forecasting 20(4), 573–587 (2004)
6. Lenk, K.: Electronic Service Delivery - A driver of public sector modernization. Journal Information Polity 7(2-3) (2002)
7. Guitouni, A., Martel, J.M.: Tentative guidelines to help choosing an appropriate MCDA method 109(2), 501–521 (September 1, 1998)
8. Laaribi, A., Chevalier, J., Martel, J.: A spatial decision aid: a multicriterion evaluation approach. Computers, Environment and Urban Systems 20(6), 351–366 (1996)
9. Hanne, T.: Intelligent Strategies for Meta Multiple Criteria Decision Making. International Series in Operations Research & Management Science. Kluwer Academic Publishers, Dordrecht (2000)
10. Kirytopoulos, K., Voulgaridou, D., Platis, A., Leopoulos, V.: An effective Markov based approach for calculating the Limit Matrix in the analytic network process. European Journal of Operational Research 214, 85–90 (2011)
11. Saaty, T.L.: Decision Making with Dependence and Feedback: The Analytic Network Process. RWS Publications, Pittsburgh (1996)
12. Figueira, J., Greco, S., EhrGott, M.: Multiple Criteria Decision Analysis: State of the Art Surveys. Springer (2005)
13. Saaty, T.L.: The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. McGraw-Hill (1980)
14. Saaty, T.L., Vargas, L.G.: The Analytic Hierarchy Process Series. University of Pittsburg, Pittsburg (1990)
15. Zahir, S.: Geometry of decision making and the vector space formulation of the analytic hierarchy process. European Journal of Operational Research 112(2), 373–396 (1999)
16. Saaty, T.L.: Decision Making in Complex Environments: the Analytic Hierarchy Process (AHP) for Decision Making and the Analytic Network Process (ANP) for Decision Making with Dependence and Feedback. RWS Publications, Pittsburgh (2003)
17. Saaty, T.L., Vargas, L.G.: Decision Making with the Analytic Network Process. Springer Science, New York (2006)
18. Weber, R., Werners, B., Zimmerman, H.J.: Planning models for research and development. European Journal of Operational Research 48, 175–188 (1990)

19. Whitaker, R.: Validation examples of the analytic hierarchy process and analytic network process. Mathematical and Computer Modelling 46(7-8), 840–859 (2007)
20. Huang, J.J., Tzeng, G.H., Ong, C.S.: Multidimensional data in multidimensional scaling using the analytic network process. Pattern Recognition Letters 26(6), 755–767 (2005)
21. Koussouris, S., Kokkinakos, P., Panopoulos, D., Askounis, D., Jain, A., Hartman, A., Zirpins, C., Georgousopoulos, C.: A Selection and Prioritisation framework for Collaborative Public Services Design. In: 5th International Conference on Methodologies, Technologies and Tools enabling e-Government, Camerino, Italy, June 30-July 1 (2011)
22. Bellos, E., Voulgaridou, D., Kirytopoulos, K., Panopoulos, D.: An MCDA Approach for Project Selection in Public Sector, Heraclion, Crete, Greece (2010)
23. Medaglia, A., Hueth, D., Mendieta, J.-C., Sefair, J.: A multiobjective model for the selection and timing for public enterprise projects. Socio-Economic Planning Sciences 42, 31–45 (2008)
24. Saaty, T.L.: Super decisions software. RWS Publications, Pittsburg (2004)
25. Bekkers, V., Zouridis, S.: Electronic Service Delivery in Public Administration: Some Trends and Issues. International Review of Administrative Sciences 65, 183–195 (1999)
26. Saaty, T.L.: Principia Mathematica Decernendi: Mathematical Principles of Decision Making. RWS Publications, PA (2010)
27. Lee, J.W., Kim, S.H.: Using analytic network process and goal programming for interdependent information system project selection. Computers & Operations Research 27, 367–382 (2000)
28. Lee, H., Kim, C., Cho, H., Park, Y.: An ANP-based technology network for identification of core technologies: A case of telecommunication technologies. Expert Systems with Applications 36(1), 894–908 (2009)
29. Jackson, P.: Public Service PerformanceEvaluation: A Strategic Perspective. Public Money & Management, 9–14 (October-December 1993)
30. Hock, L.C., Lim, J.: What Does It Mean to Optimise Public Service Delivery? Ethos (4) (April 2008)
31. Dignan, J.: What does "good enough" mean in public sector IT? OVUM (March 30, 2012), http://ovum.com
32. Sharma, A.: Evaluating public service delivery in emerging markets. In: 2011 IEEE International Conference on Service Operations, Logistics, and Informatics (SOLI), July 10-12, pp. 214–219 (2011)

E-government Benchmarking in European Union: A Multicriteria Extreme Ranking Approach

Eleftherios Siskos[1], Michail Malafekas[2], Dimitris Askounis[1], and John Psarras[1]

[1] Decision Support Systems Laboratory, School of Electrical & Computer Engineering, National Technical University of Athens, Greece
{lsiskos,askous,john}@epu.ntua.gr
[2] Decision Support Systems Laboratory, School of Electrical & Computer Engineering, National Technical University of Athens, Greece
mmalafekas@gmail.com

Abstract. E-government benchmarking is being conducted by various organizations but its assessment is based on a limited number of indicators and does not highlight the multidimensional nature of the electronically provided services. This paper outlines a multicriteria evaluation system based on four points of view: (1) infrastructures, (2) investments, (3) e-processes, and (4) users' attitude in order to evaluate European Union countries. In this paper, twenty one European Union countries are evaluated and ranked over their e-government progress. Their ranking is obtained through an additive value model which is assessed by an ordinal regression method and the use of the decision support system MIIDAS. In order to obtain robust evaluations, given the incomplete determination of inter-criteria model parameters, the extreme ranking analysis method, based on powerful mathematical programming techniques, has been applied to estimate each country's best and worst possible ranking position.

Keywords: E-government, Multicriteria analysis, Robustness, Benchmarking, Ordinal regression, Extreme ranking analysis.

1 Introduction

E-government benchmarks are used to assess the progress made by an individual country over a period of time and compare its growth against other countries. A focused assessment of e-government and other initiatives such as e-commerce, e-education, e-health, and e-science is essential if a country is to make substantial progress (Ojo et al., 2007). Benchmarks can have a significant practical impact, both political and potentially economic and can influence the development of e-government services (Bannister, 2007). The results of benchmarking and ranking studies, particularly global projects conducted by international organizations, attract considerable interest from a variety of observers, including governments (ITU, 2009).

Indices and indicators used in benchmarks are generally quantitative in nature, and collectively form a framework for assessment and ranking. Rorissa et al. (2011)

C. Douligeris et al. (Eds.): I3E 2013, IFIP AICT 399, pp. 338–348, 2013.

classified the frameworks to those based on measurable characteristics of the entities; those that use one or more subjective measures; and the rest few that employ a combination of both. They also state that frameworks based on grounded and broadly applicable measures tend to attract fewer criticisms. On the other hand, the frameworks based on subjective measures are liable to controversies and complaints, especially from those countries or institutions who consider their characterization as inaccurate. Therefore, rankings should be supported by well understood and clarified frameworks and indices as well as transparent computational procedures to maximize their acceptability by the governments and the scientific community. Two among others well-known e-government benchmarks still being conducted are those of the United Nations and the European Commission (see United Nations, 2012 and European Commission, 2010 for more details).

Recently, Siskos et al. (2012) proposed a multicriteria evaluation system based on four points of view: (1) infrastructures, (2) investments, (3) e-processes, and (4) users' attitude from which eight criteria are modeled (see below) to evaluate twenty one European Union countries for which the related data are available. Their ranking is obtained through an additive value model which is assessed by an ordinal regression method. The whole approach consists of helping decision makers (DM), experts or potential evaluators in determining their own country evaluations, based on their own value systems and their own ways of preferring, in order to propose alternative evaluation solutions in contrast to standard published benchmarks.

The aim of this paper is to extend this decision support methodology to take into account the incomplete determination of inter-criteria model parameters and to obtain a robust evaluation of the countries. These targets were achieved with the aid of the extreme ranking analysis (Kadzinski et al., 2012) that estimates each country's best and worst possible rank. This method is based on powerful mathematical programming techniques.

The paper is organized as follows: In section 2 the consistent family of criteria is briefly outlined. Section 3 presents the assessment of the multicriteria evaluation model for a single specific decision maker-evaluator while section 4 presents the extreme ranking analysis method and the obtained results. Finally section 5 concludes the paper.

2 Multicriteria Benchmarking Modeling

In order to achieve an overall assessment of global e-government, a consistent family of criteria was built according to the classical modeling methodology of Roy (1985), in the following way:

Infrastructure Criteria

g_1: Access to the web. This criterion expresses the percentage of households and businesses that have access to the web by any means.

g_2: Broadband internet connection. It shows the percentage of each country's households and businesses with a fixed broadband internet connection.

Investments Criterion

g_3: Percentage of gross domestic product (GDP) spent on information & communications technology (ICT) and research & development (R&D). ICT and R&D expenditure data were retrieved from Eurostat and the International Monetary Fund (IMF), respectively.

E-Processes Criteria

g_4: Online sophistication. It shows each country's maturity on online service delivery. The data composing this criterion stem from the European Commission's 9th Benchmark Measurement published in 2010 (Digitizing Public Services in Europe: Putting ambition into action).

g_5: E-participation. It expresses the interaction achieved between governments and citizens in a manner of information sharing, e-consultation and e-decision making. The source of the e-participation criterion is the United Nation's survey on e-government published in 2012.

Users' Attitude Criteria

g_6: Citizens' online interaction with authorities. This criterion indicates the percentage of citizens that are already using the web to interact with the authorities.

g_7. Businesses' online interaction with authorities. It indicates the percentage of businesses that are using the web to interact with the authorities.

g_8: Users' experience. This criterion expresses citizens' experience over the 20 e-services and the national portal. The data composing this criterion stem from the European Commission's 9th Benchmark Measurement.

Table 1. Criteria evaluation scales and sources

Criterion	Index	Worst level	Best level	Source
g_1	% population	0	100	Eurostat
g_2	% population	0	100	Eurostat
g_3	% GDP	0	5	Eurostat and IMF
g_4	%	0	100	European Commission
g_5	index [0-1]	0	1	United Nations
g_6	% citizens	0	100	Eurostat
g_7	% businesses	0	100	Eurostat
g_8	% index	0	100	European Commission

Table 2. Multicriteria evaluation of twenty one European countries

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
Belgium	86.5	81.5	3.3	92	0.59	28	15	68
Czech Rep.	94	90	3.3	85	0.13	35	47	22
Denmark	83.5	85.5	3.8	95	0.64	11	31	30
Germany	71.5	80.5	3.8	99	0.61	16	26	55
Estonia	86	78	4.0	97	0.69	35	18	21
Ireland	77	97	3.5	100	0.44	40	33	49
Greece	86.5	80.5	2.4	70	0.26	57	62	33
Spain	76.5	85	3.1	98	0.83	75	74	86
France	84	92	3.7	94	0.60	65	83	83
Italy	93.5	94.5	2.7	99	0.21	84	67	85
Hungary	90	80	3.4	80	0.31	57	79	80
Netherlands	75	87.5	3.5	97	0.60	80	83	80
Austria	83	78	3.3	100	0.50	85	87	90
Poland	77.5	60.5	2.7	87	0.24	88	81	81
Portugal	79	81	3.8	100	0.27	79	59.5	75.5
Slovenia	68	72.5	3.4	99	0.51	91	87	68.5
Slovakia	87	77	3.1	81	0.07	91	93	94.5
Finland	67	70	4.3	96	0.41	51	80	93.5
Sweden	79.5	65.5	4.1	99	0.49	70	88	92.5
Norway	88.5	90	2.5	92	0.50	83.5	89	87
Un. Kingdom	83.5	87.5	4,3	97	0,77	83.5	76.5	91

All details about the criteria construction techniques are thoroughly described in Siskos et al. (2012). Tables 1 and 2 present the criteria evaluation scales and the criteria scores achieved by the twenty one European countries, respectively.

3 Assessing an Overall Evaluation Model

The main target of the proposed methodological frame is the assessment of a multicriteria additive value system, for a single decision maker, that is described by the following formulae:

$$u(\boldsymbol{g}) = \sum_{i=1}^{n} p_i u_i(g_i) \tag{1}$$

$$u_i(g_{i*}) = 0,\ u_i(g_i^*) = 1, \text{ for } i = 1,2,\dots,n \tag{2}$$

$$\sum_{i=1}^{n} p_i = 1 \tag{3}$$

$$p_i \geq 0 \text{ , for } i = 1,2,\dots,n \tag{4}$$

where $\boldsymbol{g} = (g_1, g_2, \ldots, g_n)$ is the performance vector of a country on the n criteria; g_{i*} and g_i^* are the least and most preferable levels of the criterion g_i, respectively; $u_i(g_i)$, $i = 1,2,\ldots,n$ are non decreasing marginal value functions of the performances g_i, $i = 1,2,\ldots,n$; and p_i is the relative weight of the i-th function $u_i(g_i)$. Thus, for a given country a, $\boldsymbol{g}(a)$ and $u[\boldsymbol{g}(a)]$ represent the multicriteria vector of performances and the global value of the country a respectively.

Both the marginal and the global value functions have the monotonicity property of the true criterion. For instance, in the case of the global value function, given two countries a and b the following properties hold:

$$u[\boldsymbol{g}(a)] > u[\boldsymbol{g}(b)] \Leftrightarrow a\ P\ b \quad \text{(Preference)} \tag{5}$$

$$u[\boldsymbol{g}(a)] = u[\boldsymbol{g}(b)] \Leftrightarrow a\ I\ b \quad \text{(Indifference)} \tag{6}$$

The necessary hypothesis to validate an additive value function for a given decision maker (DM) is the preferential independence of all the criteria (see Keenney and Raiffa, 1976, Keeney, 1992 for instance). A pair of criteria (g_i , g_j) is preferentially independent from the rest of the criteria when the trade-offs between the g_i and g_j criteria are not dependent on the values of the rest of the criteria. All criteria are supposed to be preferentially independent when the same condition holds for all pairs of criteria. When the u_i functions in formula (1) are already assessed, the linear model (1)-(4) exists if and only if the inter-criteria parameters (weighs) p_i are constant substitution rates (value trade-offs) between u_i.

This value system can be obtained utilizing various methods (see Keeney and Raiffa, 1976, Keeney, 1992, Figueira et al., 2005). Because of the objective difficulties to convince decision makers in externalizing tradeoffs between heterogeneous criteria and verify the preferential conditions cited above, decision analysts usually prefer to infer the DM's additive value function from global preference structures, by applying disaggregation or ordinal regression methods (see Jacquet-Lagrèze and Siskos, 1982, 2001, Greco et al., 2008, 2010).

In this study the disaggregation UTA II method is implemented by assessing the additive model (1)-(4) in two phases: In the first phase the expert is asked to assign some value points $u_i(g_i)$ of the corresponding evaluation scale for every criterion separately. Then, each marginal value function is optimally fitted (see Fig. 1) and accepted by the DM. In a second phase, the criteria weights p_i are estimated using inference procedures (see next section).

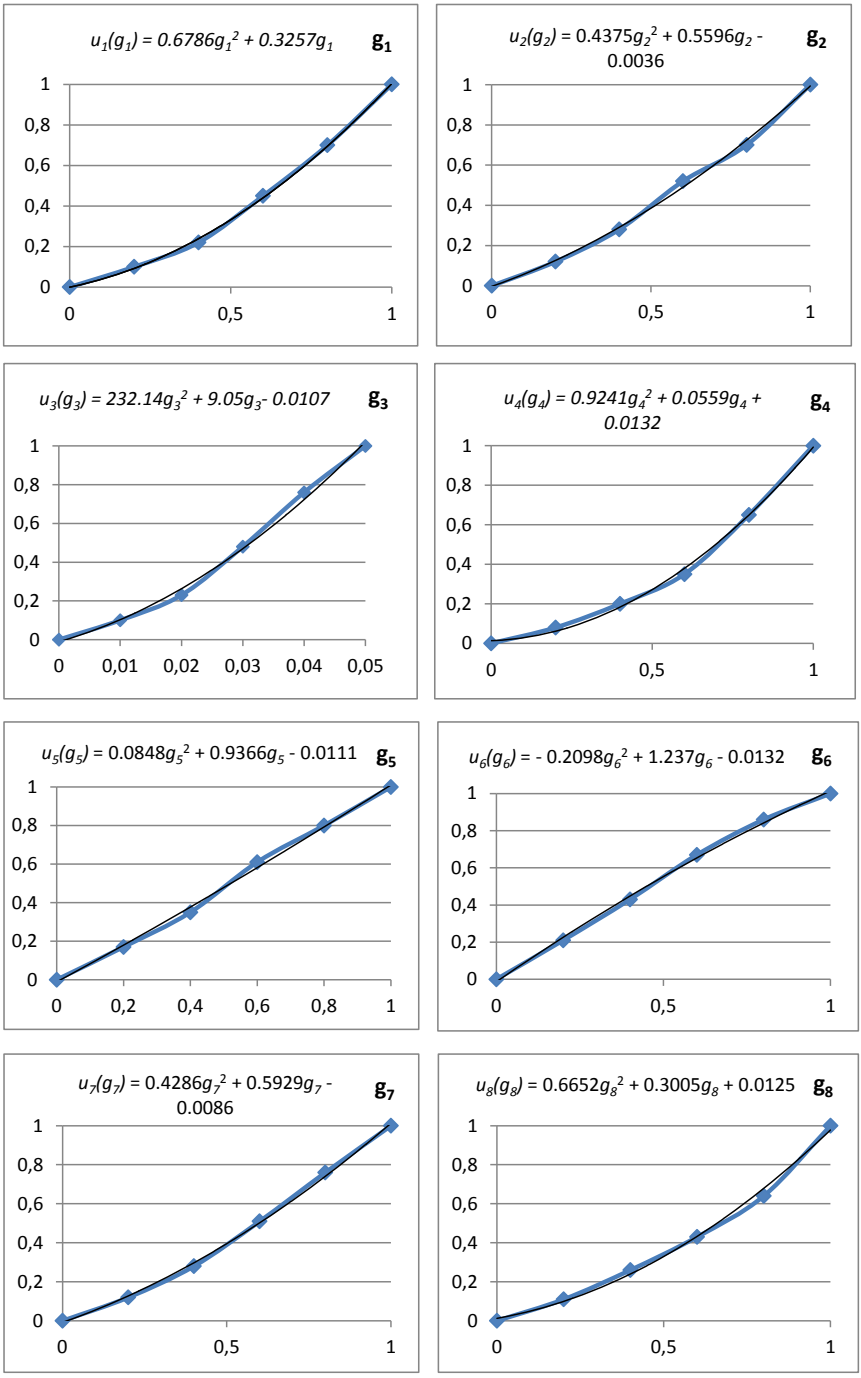


Fig. 1. Marginal value functions of the eight criteria

4 Estimation of Weights and Extreme Ranking Analysis

Through this phase the UTA II disaggregation procedure is used to infer the inter-criteria parameters p_i, $i = 1,2, \dots, n$. Specifically, the DM-evaluator is asked to give a ranking (weak order) on a set of reference countries $A_r = (a_1, a_2, \dots, a_k)$, that are fictitious country profiles differing on two or at most three criteria values. The reference countries are ranked by the DM in such a way that a_1 is the head and a_k the tail of the ranking.

Therefore, for every pair of consecutive countries (a_m, a_{m+1}) holds, either $a_m P a_{m+1}$ (preference of a_m) or $a_m I a_{m+1}$ (indifference). UTA II solves the linear program (7)-(11) below that has k constraints because of the transitivity of the (P, I) preference system.

$$[min]F,\ F = \sum_{i=1}^{k}\left(\sigma^{+}(\alpha_i) + \sigma^{-}(\alpha_i)\right) \tag{7}$$

Subject to:

for $m = 1, 2, \dots, k-1$

$$\sum_{i=1}^{n} p_i\, u_i[g_i(a_m)] - \sigma^{+}(\alpha_m) + \sigma^{-}(\alpha_m) - [\sum_{i=1}^{n} p_i\, u_i[g_i(a_{m+1})] - \sigma^{+}(a_{m+1}) + \sigma^{-}(a_{m+1})] \geq \delta\ ,\ \text{if } a_m\ P\ a_{m+1} \tag{8}$$

or

$$= 0 \quad \text{if } a_m\ I\ a_{m+1} \tag{9}$$

$$\sum_{i=1}^{n} p_i = 1 \tag{10}$$

$$p_i \geq 0, for\ \ i = 1, 2, \dots, n;\ \sigma^{+}(a_j) \geq 0, \sigma^{-}(a_j) \geq 0, for\ j = 2, \dots, k \tag{11}$$

where δ is a small positive number, equal to 0.001 for instance; $g_i(a_m)$ is the evaluation of the a_m country on the i-th criterion and $u_i[g_i(a_m)]$ its corresponding marginal value; $\sigma^{+}(a_j)$, $\sigma^{-}(a_j)$ are the over-estimation and the under-estimation errors concerning the j-th country's position, respectively.

This technique was applied for a set of thirteen country profiles and a zero error sum was obtained ($F = 0$). The optimally most characteristic weighting factors are reported in formula (12) while the corresponding ranking of the European countries is presented in Table 3. More details about UTA II illustration is given in Siskos et al. (2012).

$$u(\boldsymbol{g}) = 0.1276u_1(g_1) + 0.1607u_2(g_2) + 0.1097u_3(g_3) + 0.2579u_4(g_4) + 0.0743u_5(g_5) + 0.1209u_6(g_6) + 0.0536u_7(g_7) + 0.0952u_8(g_8) \tag{12}$$

$\boldsymbol{g} = (g_1, g_2, \dots, g_8)$, is the performance vector of a country on the eight criteria.

Table 3. E-government ranking of the twenty one European countries

Rank position	European Country	Global Value
1	Sweden	0.825
2	Denmark	0.821
3-4	Finland	0.796
3-4	Netherlands	0.796
5	Norway	0.765
6-7	Germany	0.745
6-7	United Kingdom	0.744
8	France	0.738
9	Estonia	0.729
10-11	Austria	0.701
10-11	Slovenia	0.701
12	Spain	0.693
13	Belgium	0.686
14	Ireland	0.679
15	Portugal	0.633
16	Czech Republic	0.582
17	Slovakia	0.578
18	Hungary	0.568
19	Poland	0.548
20	Italy	0.533
21	Greece	0.467

However, the above estimation procedure bears robustness issues. In fact, there exists an infinite number of weighting vectors that are optimally consistent with the whole set of constraints (8)-(11). In order to study the impact of this indetermination on the ranking of the countries the extreme ranking analysis of Kadzinski et al. (2012) has been applied with the aid of the GAMS platform.

The extreme ranking analysis algorithm examines each country individually and estimates the best and worst possible rank it can achieve. The methodology leading to the estimation of the best possible rank of each country is outlined below.

In order to determine the best possible rank of a country A, taking into consideration all the possible combinations of the criteria weighting factors, the number N_A^* of the countries that surpass country A in the ranking under any

circumstances is calculated. The countries that surpass country A in the ranking for a limited number of combinations of the criteria weights are not included in the N_A^* set. Therefore, the best possible ranking position that can be achieved by the country A is N_A^*+1.

Thus, the problem is reduced to the calculation of the N_A^* set for each individual country. This set is calculated through the modeling and the solution of the mixed integer programming problem presented below:

$$[min]F = \sum_{b\in A\setminus\{a\}} u_b \tag{13}$$

Subject to:

$$\text{Constraints (8) – (11)} \tag{14}$$

$$U(\alpha) \geq U(b) - Mu_b, \forall\ b \in A\setminus\{\alpha\} \tag{15}$$

where M is an auxiliary variable equal to a big positive value, and u_b is a binary variable associated with comparison of the country *A* to another country *B*. There exist $N - 1$ such variables, each corresponding to $b \in A\setminus\{\alpha\}$. N is the total number of the countries under evaluation, i.e. 21.

The determination of the worst possible ranking of a country *A*, requires a similar procedure. In this case, however, it is estimated the number of countries N_{A^*} that achieve a worse ranking position for all possible combinations of the weighting factors. Therefore, the worst possible rank a country *A* can achieve is $N - N_{A^*}$.

N_{A^*} set is calculated through the solution of the integer programming problem outlined below:

$$[min]F = \sum_{b\in A\setminus\{a\}} u_b \tag{16}$$

Subject to:

$$\text{Constraints (8) - (11)} \tag{17}$$

$$U(b) \geq U(\alpha) - Mu_b, \forall\ b \in A\setminus\{\alpha\} \tag{18}$$

The extreme ranking positions of the 21 European countries assessed, are graphically presented in Fig. 2.

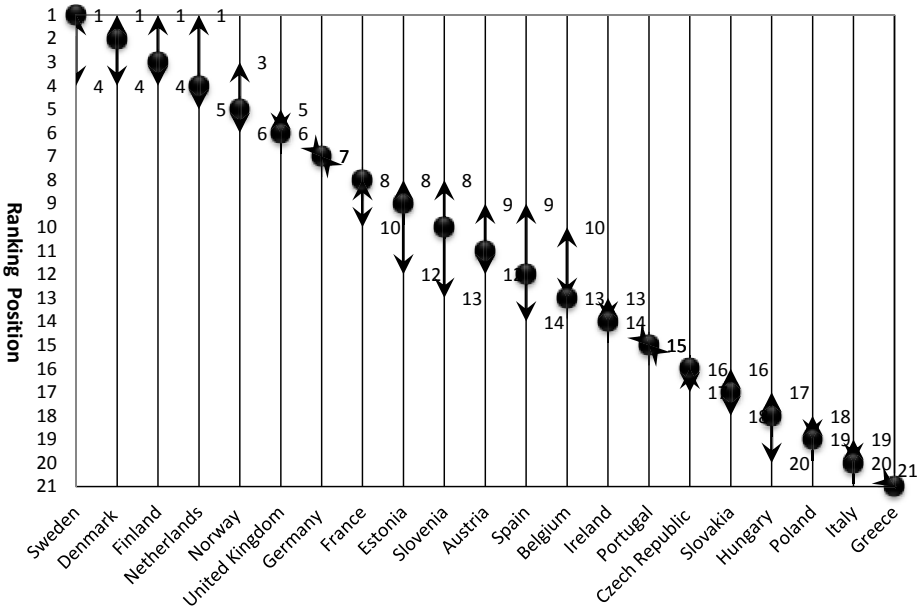


Fig. 2. Extreme ranking positions of twenty one European countries

5 Conclusion

The paper addressed the robust assessment of global e-government based on multiple criteria and special extreme ranking procedures. The proposed approach focused on the evaluation of European countries according to the standards of a benchmark.

The e-government evaluation process is an independent procedure enabling each individual to specify his (her) own preferences on criteria value functions and weights, and results in a personalized ranking of the countries. In other words, each evaluator has control over his (her) set of criteria and the assessment of his (her) own evaluation model. The proposed multicriteria techniques offer the possibility to combine different preferences and considerations of multiple decision makers and merge them easily through interactive iterative processes.

The next research steps include the development of robustness control procedures based on cardinal and visualization measures as well as the development of a decision support system aiding anyone to form his own e-government benchmarking.

Acknowledgements. This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALES. Investing in knowledge society through the European Social Fund.

References

1. Bannister, F.: The curse of the benchmark: an assessment of the validity and value of e-government comparisons. International Review of Administrative Sciences 73(2), 171–188 (2007)
2. European Commission: Digitizing public services in Europe: Putting ambition into action, 9th benchmark measurement, Directorate General Information Society and Media, Brussels (2010)
3. Figueira, J., Greco, S., Ehrgott, M.: State-of-Art of Multiple Criteria Decision Analysis. Kluwer Academic Publishers, Dortrecht (2005)
4. Greco, S., Mousseau, V., Słowiński, R.: Ordinal regression revisited: Multiple criteria ranking using a set of additive value functions. European Journal of Operational Research 191(2), 416–436 (2008)
5. Greco, S., Słowiński, R., Figueira, J., Mousseau, V.: Robust ordinal regression. In: Ehrgott, M., Figueira, J., Greco, S. (eds.) Trends in Multiple Criteria Decision Analysis, ch. 9, pp. 241–283. Springer, New York (2010)
6. ITU: Measuring the information society: The ICT development index (2009), http://www.itu.int/ITU-D/ict/publications/idi/2009/material/IDI2009_w5.pdf (retrieved November 10, 2012)
7. Jacquet-Lagrèze, E., Siskos, J.: Assessing a set of additive utility functions for multicriteria decision making. European Journal of Operational Research 10(2), 151–164 (1982)
8. Jacquet-Lagrèze, E., Siskos, J.: Preference disaggregation: 20 years of MCDA experience. European Journal of Operational Research 130, 233–245 (2001)
9. Kadzinski, M., Greco, S., Slowinski, R.: Extreme ranking analysis in robust ordinal regression. Omega 40, 488–501 (2012)
10. Keeney, R.L.: Value-focused thinking: A path to creative decision making. Harvard U.P., London (1992)
11. Keeney, R.L., Raiffa, H.: Decisions with multiple objectives: Preferences and value trade-offs. John Wiley and Sons, New York (1976)
12. Ojo, A., Janowski, T., Estevez, E.: Determining progress towards e-Government - What are the core indicators? (2007), http://www.iist.unu.edu/newrh/III/1/docs/techreports/report360.pdf (retrieved November 10, 2012)
13. Rorissa, A., Demissie, D., Pardo, T.: Benchmarking e-Government: A comparison of frameworks for computing e-Government index and ranking. Government Information Quarterly 28, 354–362 (2011)
14. Roy, B.: Méthodologie multicritère d'aide à la décision. Economica, Paris (1985)
15. Siskos, E., Askounis, D., Psarras, J.: Multicriteria decision support for global e-government evaluation, Working paper, Decision Support Systems Laboratory, School of Electrical & Computer Engineering. National Technical University of Athens, Greece (2012)
16. United Nations: E-government survey 2012, E-government for the people (2012), http://www2.unpan.org/egovkb/global_reports/12report.htm (retrieved November 10, 2012)

Author Index