SOTICS 2012

The Second International Conference on Social Eco-Informatics

ISBN: 978-1-61208-228-8

October 21-26, 2012

Venice, Italy

SOTICS 2012 Editors

Lasse Berntzen, Vestfold University College - Tønsberg, Norwa

Petre Dini, Concordia University, Canada & China Space Agency, China

SOTICS 2012

Foreword

The Second International Conference on Social Eco-Informatics [SOTICS 2012], held between October 21-26, 2012 in Venice, Italy, continued a series of events on social eco-informatics, bridging different social and informatics concepts by considering digital domains, social metrics, social applications, services, and challenges. Academic and industrial contributions were expected on algorithms, mechanisms, models, services dealing with challenges in social eco-systems.

The systems comprising human and information features form a complex mix of social sciences and informatics concepts embraced by the so-called social eco-systems. These are interdisciplinary approaches on social phenomena supported by advanced informatics solutions. It is quit intriguing that the impact on society is little studied despite a few experiments. Recently, also Google was labeled as a company that does not contribute to brain development by instantly showing the response for a query. This is in contrast to the fact that it has been proven that not showing the definitive answer directly facilitates a learning process better. Also, studies show that e-book reading takes more times than reading a printed one. Digital libraries and deep web offer a vast spectrum of information. Large scale digital library and access-free digital libraries, as well as social networks and tools constitute challenges in terms of accessibility, trust, privacy, and user satisfaction. The current questions concern the trade-off, where our actions must focus, and how to increase the accessibility to eSocial resources.

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

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the SOTICS 2012 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that SOTICS 2012 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in advanced social eco-informatics.

We are convinced that the participants found the event useful and communications very open. We also hope the attendees enjoyed the charm of Venice, Italy.

SOTICS Chairs:

Philip Davis, Bournemouth and Poole College - Bournemouth, UK
David Newell, Bournemouth University - Bournemouth, UK

Petre Dini, Concordia University, Canada & IARIA, USA
Magdalini Eirinaki, San Jose State University, USA
Krzysztof Juszczyszyn, Wrocław University of Technology, Poland
Ulrike Lucke, University of Potsdam, Germany
Abdulrahman Yarali, Murray State University, USA
Feng Gao, Microsoft Corporation, USA
Lynne Hall, University of Sunderland, UK
Jennifer Watkins, Los Alamos National Laboratory, USA
Tom Erickson, IBM T. J. Watson Research Center, USA
Paolo Garza, Politecnico di Milano, Italy
Claus Atzenbeck, Hof University, Germany / Aalborg University, Denmark
Nima Dokoohaki, Royal Institute of Technology (KTH) - Stockholm, Sweden
Christine Langeron, Hubert Curien Laboratory, Lyon University, France
Sandra Sendra Compte, Polytechnic University of Valencia, Spain

SOTICS 2012

Committee

SOTICS General Chairs

Philip Davis, Bournemouth and Poole College - Bournemouth, UK
David Newell, Bournemouth University - Bournemouth, UK

SOTICS Advisory Chairs

Petre Dini, Concordia University, Canada & IARIA, USA
Magdalini Eirinaki, San Jose State University, USA
Krzysztof Juszczyszyn, Wrocław University of Technology, Poland
Ulrike Lucke, University of Potsdam, Germany
Abdulrahman Yarali, Murray State University, USA

SOTICS Special Area Chairs on Social Networks

Feng Gao, Microsoft Corporation, USA
Lynne Hall, University of Sunderland, UK

SOTICS Special Area Chairs on eGovernment

Jennifer Watkins, Los Alamos National Laboratory, USA

SOTICS Special Area Chairs on Digital Cities

Tom Erickson, IBM T. J. Watson Research Center, USA

SOTICS Special Area Chairs on Digital Libraries

Paolo Garza, Politecnico di Milano, Italy

SOTICS Special Area Chairs on Media

Claus Atzenbeck, Hof University, Germany / Aalborg University, Denmark

SOTICS Publicity Chairs

Nima Dokoohaki, Royal Institute of Technology (KTH) - Stockholm, Sweden
Christine Langeron, Hubert Curien Laboratory, Lyon University, France
Sandra Sendra Compte, Polytechnic University of Valencia, Spain

SOTICS 2012 Technical Program Committee

Gil Ad Ariely, California State University - Chico | (IDC) Herzliya, Israel
Mehdi Asgarkhani, CPIT – Christchurch, New Zealand
Lasse Berntzen, Vestfold University College, Norway
Anne Boyer, Université de Lorraine, France
Piotr Bródka, Wroclaw University of Technology, Poland
Erik Buchmann, Karlsruhe Institute of Technology (KIT), Germnay
María Luisa Carrió-Pastor, Universitat Politècnica de València, Spain
Aspassia Daskalopulu, University of Thessaly, Greece
Guillermo De Ita, Universidad Politécnica de Puebla, Mexico
Petra Deger, PH Heidelberg, Germany
Matteo Dell'Amico, EURECOM, France
Nima Dokoohaki, Royal Institute of Technology (KTH) - Stockholm, Sweden
Prokopios Drogkaris,University of the Aegean - Karlovasi, Greece
Magdalini Eirinaki, San Jose State University, USA
Thomas Erickson, IBM's Watson Labs, USA
Larbi Esmahi, Athabasca University, Canada
M. Caremen Fernandez-Gago, University of Malaga, Spain
Jean Vincent Fonou-Dombeu, Vaal University of Technology, South Africa
Jean-Gabriel Ganascia, University Pierre and Marie Curie, France
Paolo Garza, Politecnico di Milano, Italy
Rimantas Gatautis, Kaunas University of Technology, Lithuania
Chris Geiger, Duesseldorf University of Applied Sciences, Germany
Lilia Georgieva, Heriot-Watt University, UK
William I. Grosky, University of Michigan-Dearborn, USA
Richard Gunstone, Bournemouth University, UK
Panos Hahamis, University of Westminster – London, UK
Lynne Hall, University of Sunderland, UK
R. Gy Hashim, Universiti Teknologi MARA (UiTM) - Shah Alam, Malaysia
Yun Huang, Northwestern University, USA
Omar S. Hujran, Princess Sumaya University for Technology – Amman, Jordan
Colette Hume, University of Sunderland, UK
Carlos E. Jiménez, IEEE Technology Management Council Spain, Spain
Maria João Simões, Universidade da Beira Interior - Covilhã, Portugal
Georgios Kapogiannis, University of Salford, UK
Przemyslaw Kazienko, Wroclaw University of Technology, Poland
Babak Khosravifar, Concordia University, Canada
Christian Kittl, evolaris next level GmbH, Germany
Tewart Kowalski, Stockholm University, Sweden
Renaud Lambiotte, FUNDP - Namur, Belgium
Christine Largeron, Lyon University, France
Shou-de Lin, National Taiwan University, Taiwan
Xumin Liu, Rochester Institute of Technology, USA
Ulrike Lucke, University of Potsdam, Germany
Paul Lukowicz, German Research Center for Artificial Intelligence (DFKI) - Kaiserslautern, Germany
Lorenzo Magnani, University of Pavia, Italy / Sun Yat-sen University, Guangzhou (Canton), P. R. China
Philippe Mathieu, Université Lille 1, France

Copyright Information

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

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

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

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

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

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

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

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

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

Table of Contents

Ethics and Social Eco-Informatics

Richard Lucas
Information Systems
University of Canberra
Canberra, Australia
and
Centre for Applied Philosophy and Public Ethics Charles Sturt University
Email: Richard.Lucas@canberra.edu.au

***Abstract*—Ethics problems in emerging fields, typically, do not get advance consideration. It is not until after an ethical lapse occurs in a field of study does ethics get a serious examination. Social Eco-informatics is no exception. I will show the extent of the existing consideration in the academic literature concerning eco-informatics of ethics problems and their ethical basis. Lessons from the history of emerging fields in information technologies will be examined as a way of pointing out the detail of ethical problems that this field might encounter. This paper looks at the ethics-first – ethics-last debate and offers some guidance for researchers in this emerging field. The challenge for emerging fields such as eco-informatics lies in keeping ahead of potential ethics problems not reacting to them after they occur. Finally, I use a paper on social eco-informatics to point to possible places where the eco-informatic researcher might look to include ethical considerations.**

***Index Terms*—ethics first; ethics last; eco-informatics; personal ethics;**

It needs no proof that neither economic activities nor any other class of human activities can rightly be made independent of moral laws. (Keynes, 1890, 24) [8]

I. Introduction

I would have liked to have written a paper about ethical problems in social eco-informatics, showing significant and deep insight into common concerns such as privacy protection. However when I began to examine the usual ethical suspects, harm, rights violations, unprofessionalism, and so on I became aware of the lack of extant research into ethical problems in the field. I had encountered this problem before when examining the moral issues in enterprise computing [11]: No literature. I could have retreated to the established and field of environmental ethics and expounded on, for example, privacy protection and what might be done about it, but my interest was piqued: I wondered why there might be nothing so far? Was it just a coincidence that two emerging IT fields have no current, or past, interest in their ethical dimensions? Out of this wonder emerged this paper.

A. A word about terms

The usual practice in science research is to precisely define terms. For this paper, fortunately, such a requirement is unnecessary when discussing ethics and social eco-informatics. Why is this?

Even if there were precise definitions of the individual terms ETHICS and SOCIAL ECO-INFORMATICS that all could agree on they, as discussed in the next part, would have no effect on the thrust of this paper. What I was looking for in this paper was the *occurrence* of the term pair ETHICS and SOCIAL ECO-INFORMATICS, regardless of what the individual terms might mean to any particular individual.

Nonetheless, ETHICS is a sufficiently common term that, while there may be some disagreement about what counts as ethical (or more usually, unethical), there is no substantive disagreement that everyone has a conception of what counts as being ethical. A common example might be:

Ethics is a set of values for determining right and wrong.

I could find no such common example for SOCIAL ECO-INFORMATICS. For this paper I take the characterization of eco-informatics used by Kineman and Kumar [9].

> represent natural biodiversity and ecosystem phenomena, and to communicate such information to society, for science, valuation, management, and policy

II. What is there to be ethical about with Eco-informatics?

Apparently nothing. Ethics, it seems, is not on the radar of researchers into eco-informatics.

In a search of Google Scholar conducted at the last possible moment for this paper (18 June 2012) using the keyword phrases “eco-informatics” + “ethics”I searched for the occurrence of research papers on the topic of ethical problems in eco-informatics. I had originally planned to take the results of this search and refine it by adding “social”. However, as the results below show this became unnecessary. Also note that a mere seven search results were in the “ethical”list that did not also appear in the “ethics”list: this made discussing them separately unnecessary.

A. The Google Scholar Search

Google Scholar was searched for academic papers and books that examined the phrase “eco-informatics”. The term eco-informatics alone without additional search terms resulted in a mere 303 results. The search uncovered the first occurrence of the term “eco-informatics”combined with “ethics”and

 ISBN: 978-1-61208-228-8

"ethical"as part of the more weildy "geo-eco-informatics"(or geographical and ecological information systems) in Koshkariov, Tikunov, and Trofimov (1989) [10]. This makes the use of the term eco-informatics a relatively new phenomenon. The results were then categorized into nine groups:

TABLE I
KEYWORD SEARCH RESULTS CATEGORIES "ECO-INFORMATICS" + "ETHICS"OR "ETHICAL"

	Category	ethics	ethical
1	Jnl or Org Code of Ethics	1	0
2	Another field's ethics	5	4
3	Reference to an ethics paper	3	2
4	Ethics is important	5	2
5	Personal Ethics	1	1
6	Soft Science	1	1
7	Actual Examples	0	0
8	Moral evaluation	0	0
9	Unable to locate term	39	9
	Categories Totals	55	19

Now, I do not want to make too much of these results. I am fully aware that the ethics of eco-informatics could be regarded as a melding of the fields of environmental ethics and information ethics and merely examining each of these fields ought to be sufficient to identify any ethical problems and be aware of any moral evaluations. This may be true but any connections to these already established fields does not make the point that I want to make here. The point here is that if a field chooses to create a new term, such as eco-informatics, then the field, presumably, has substantial reasons to distinguish itself from other closely related fields. This difference then ought to result in unique ethical problems and present unique opportunities to show its moral worth.

A cursory analysis of this shows some interesting features.

1) Jnl or Org Code of Ethics: This category refers to those search results that mention a journal or organizational code of ethics. One search result made reference to the need to follow a code of ethics but did not say why or what the relevance of such a code might be.

2) Another field's ethics: This category refers to those search results that mention another field's ethics, environmental ethics, for example. Several search results (5 and 4) made reference to another academic field's ethics. Environmental and information ethics both got a mention. However no details were drawn out nor any ethical problems relating specifically to eco-informatics were mentioned.

3) Reference to an ethics paper: This category refers to those search results that specifically mention a paper that has ethics in its title or is in an established applied ethics journal. Three search results included a reference to a paper in an ethics journal such as *Journal of Information Ethics* and *Journal of Environmental Ethics*. However no use was made by the articles found to discuss the ethics of eco-informatics.

4) Ethics is important: This category refers to those search results that, in some form or other, emphasise the importance of ethics or ethical concepts (such as harm, benefit,trust, autonomy, honesty, virtue). Five search results mentioned the importance of ethics but did not say how or why.

5) Personal Ethics: This category refers to those search results that refer to personal or a person's ethics. One search result said that personal ethics were important when dealing with eco-informatics. Again there was no caching out of this claim nor evidence of how this might be.

6) Soft Science: This category refers to those search results that refer to ethics obliquely by bundling it in with other social or 'soft' sciences. For one search result the authors, referring to those constructing eco-informatics systems, wrote that it was important to take into account soft sciences such as ethics, management, and society into account. However, "Soft or social sciences, in so far as they embrace environmental issues, history, ethics and philosophy have til now run a very poor second." [14]

7) Actual Examples: This category refers to those search results that refer to specific ethical problems that have occurred, and are named in connection with eco-informatics. No search result gave actual examples of ethical problems. I would have also taken (and did search for), the usually erroneously named, ethical dilemma but there were none. This will be one of the two main discussion points in the next subsection.

8) Moral Evaluation: This category refers to those search results that mention any sort of evaluation of eco-informatics, eco-informatics systems in terms of its moral or social worth. No search result gave actual examples of ethical problems. I would have included any indication of the ethical merit of specific examples of eco-informatics (data, information, or systems) , the usually erroneously named, ethical dilemma but there were none. This will be one of the two main discussion points in the next subsection.

B. Some (non-scientific) analysis

My conclusion is that ethical eco-informatics, as a serious field of study, has been much neglected.

I do not intend to provide a sophisticated statistical analysis of the above search nor an in-depth defence of my conclusion. That is not my goal here. My goal here is to point out that for a term that is 23 years (see [10]) old there are no serious academic papers tackling the idea that there might be ethical problems lurking in the background or that the moral worth of eco-informatics be examined. This might be explained by the fact that no ethical problems have occurred in the field; that the moral worth of eco-informatics is so obvious so as to not need stating.

Perhaps, the findings above are just artefacts of history, that will go the way of all technological implementations of informations systems: we will work out the problems eventually. Is it merely a matter of time before researches into eco-informatics and ethics come good; that academics have a go at examining the ethical problems (potential and actual) to measure the ethical merit?

Put another way, is eco-informatics any different to any other field? After all, it seems that ethics gets mentioned, eventually, in every field. However, if you look closely at nearly any field it turns out that ethics gets a mention only

 ISBN: 978-1-61208-228-8

after some unethical, usually catastrophic, event has occurred. That is, ethics gets considered last.

III. Ethics first or ethics last?

It seems clear to me that in the emerging field encompassing the environment, technology, and information (eco-informatics) there has been no forethought about its ethical implications. That is, ethics has not come first. Also pointed out previously is that this is nothing new, it happens all the time. But why should this be so? Why does ethics not get considered first, as the field emerges? Why is ethics not discussed ahead of time?

A. Lessons from history

To frame a discussion on whether ethics ought to be done first or last I will use, one traditional area of information and communication technology (ICT) ethics, systems development, and two emerging areas, radio frequency identification device (RFID) tags and data matching. In a forthcoming book, *Professionalism in the Information and Communication Technology Industry* [20] Bernd Stahl and Simon Rogerson, say that: "A central problem of the ethics of technology is that it [ethics] tends to arrive too late". Michael Kirby the recently retired (2009) Justice of the High Court of Australia, showed that there were many difficulties in regulating new technologies, ICT specially. In a similar vein, Philippe Goujon and Catherine Flick state that "the strong push for technology development too often obscures the need for any deep ethical consideration before a technical project is funded, developed and deployed".

Researchers can, either react to the technology, that is the ethics is done last, after the technology has been developed, or, it can be proactive, that is be done first, before the technology is developed. This ethics first – ethics last, approach seems to raise what is commonly known as the Collingridge [2] dilemma. If the ethics is done before the development has impacts it is difficult to predict what those impacts will be and if done after development it is difficult to control these impacts. Croy [3] formulates the dilemma this way:

> Either a technology is in a relatively early stage of development when it is unknown what changes should be made, or a technology is in a relatively late stage of development when change is expensive, difficult and time-consuming.
>
> If the former, then control is not possible.
>
> If the latter, then control is not feasible.
>
> Therefore, either controlling technology is not possible, or controlling technology is not feasible.

In short, controlling technology is not possible because prediction is so unreliable and is not feasible because, once the technology is developed; change is difficult. While this appears to pose real difficulties, it is not quite the dilemma that it initially appears to be. Some prediction can be informative and some control after the development is possible. Furthermore, it is possible, to some extent, to slip between the horns.

B. What is the difference between reactive and proactive active ethics?

1) Reactive ethics: I will show the difference between reactive and proactive ethics through a brief look at the history of the closely related (to the informatics part of eco-informatics) field of the ethics of information technology. Problems in IT ethics are not different or new in the sense that they are different in kind from other ethical problems. Moral philosophy has been studied , at least, for the past 2500 years, and the ethical issues in computing are an extension of this. What is new and different is that the development and use of computers has raised old questions in interestingly new and different ways [7], often creating what Moor [12] calls policy vacuums: problems without worked out ways of dealing with them. The work of the computer ethicists then is to develop policies to fill those vacuums. One example in computing is hacking. Breaking into someone's computer account is in some ways like breaking into someone's house, but there are interesting differences. It is a logical rather than a physical entering. Another is unauthorised copying of software, which is a bit like unauthorised copying of a book, and a bit like taking a television set, but there are significant differences. There are also questions relating to work and the loss or creation of skills, which arise in a unique way. This way of doing computer ethics is essentially reactive [7] and, unfortunately, the most common role of applied ethics in general.

2) Proactive ethics: When a proactive, or ethics-first, approach is taken, the emphasis is different. One is much more likely, and in fact it is necessary, to think carefully about what is wanted from the technology, and that involves thinking about what sort of life one thinks is a good one. This approach means taking action that will guide the development of the technology in a particular manner. This proactive stance also highlights a more positive view of applied ethics. Ethicists are frequently seen as playing only a negative role, always criticising and attempting to hinder development. While this view is to some extent correct, it is not the only thing that ethicists should be doing. Technology clearly has a positive role. In many ways our lives are much better because of various technologies.

In software engineering (a common area of ICT), calls for ethics to be considered early in the development process are not new; Don Gotterbarn [6] has been promoting a piece of software called SODIS (for the ethical evaluation of software development projects) for years. This, however, is rare. In ICT development, as in most technical development, ethical questions have generally been left until problems appear.

3) Both: So, should ITC ethics be reactive, proactive, or both? That is, should ICT ethics just respond to existing problems, try to anticipate problems, or both respond and anticipate?

The argument here is that the ethics first model and the ethics last model are popular but poor solutions to a false dichotomy (see Weckert [19] for a detailed examination). ICT ethics is not something one can complete satisfactorily either

 ISBN: 978-1-61208-228-8

first or last but something that needs be done continually as the technology develops and as its potential consequences become better understood. Ethics is dynamic in that the factual component on which it relies has to be continually updated. Norbert Wiener outlined this approach in 1960. In talking about automated machines, he writes:

> To be effective in warding off disastrous consequences, our understanding of our man-made machines should in general develop *pari passu* [in step with] with the performance of the machine.
> (Wiener, 1960, 1355) [22]

What he is suggesting is that if we wait until the technology has been developed, it may be too late to avoid disastrous consequences. Predicting is hazardous and he does not suggest that these machines should not be developed on the grounds that they may produce undesirable consequences. Rather, the understanding must develop in step with the development of the technology's systems. Similarly, the contention here is that the understanding of ethical questions must develop as the technology develops. This will be dynamic; partly reactive, partly *pari passu*, and partly proactive; continually returning to the technology to understanding how it is developing and what its actual or likely consequences are.

Technologists must work with ethicists to predict future problem areas. This is of course fraught with danger given the uncertain nature of prediction, but if care is taken, it is a useful and important activity. In many cases it is at least plausible that ethical problems will arise and that we can know something about what they will be like. Take the case of RFID chips (These are the chips that are embedded in clothing or otherwise attached to desirable and disposable goods. They emit a radio frequency which is detected by a scanning device within its range.). It is known that they have been developed significantly compared to the passive chips currently used in libraries and other places and research is continuing. For example a new passive RFID chip is under development that will be able to be printed onto clothing, and paper, with a range eventually of around 300 metres (Williams [21]).

As another example, think of data matching (the process of aggregating information across a number of disparate sources. Usually, these, many sources were never intended to be put together or matched) within a large organisation. Consider the enterprise as a multinational, multi-industry conglomeration. This enterprise has a number of customer loyalty cards. People who took out the individual, separate, loyalty cards never imagined that the information they contain would be combined to provide a comprehensive profile of the card holder. Neither did the designers of the loyalty systems. Were it only restricted to the giving bonuses for loyalty but it never ends there. It gets extended to more and more activities that are less and less relevant to the card holder.

It is, now, fairly clear that developments in ICT in this context, will lead to further consolidation and aggregation of citizens' information. Given these developments and uses of the citizen data already in other contexts (Thornley, et al. [17]), it is certainly plausible that government agencies and private enterprise (for example the marketing departments for both groups) will extend their uses in ways that are increasingly threatening to privacy and make them much more than mere isolated databases. This is not inevitable of course: I am not a proponent of technological determinism, the view that technology will move on regardless of what we do. However, given the extent to which the values of efficiency and productivity seem to override other values in the assessment of technologies and their uses, the pressures to extent their uses will be intense if it is seen to aid efficiency. Resistance might also be weak. In other areas privacy is eroded a little at a time in a way that is almost unnoticeable. Imagine if, say about twenty years ago before paying for groceries with credit cards was popular, the supermarkets had employed people to sit and watch all customers as they paid, and made a record of their purchases, names, and addresses. Most of us probably would have been a little concerned about this. Now it all happens automatically for those who pay by credit card; most people do not give it a second thought. This suggests that what the eye doesnt see doesnt really bother most people: what happens gradually goes largely unnoticed.

C. So What?

It is here that proactive ethics comes in to play.

1) RFIDs: The argument is not that RFID chips should not be used in libraries because of future dangers. The question to be asked is, is it morally responsible to use the chips in certain extended ways? And how can the technology be designed to mitigate the dangers? What kind of regulations need to be put in place to control the use of the chips? These are all legitimate questions for the proactive ethicist and their answers should feed into policy decisions about the future use of RFID chips.

2) Data Matching: In the other example of data matching the questions to be asked are; is it morally responsible to use loyalty card information in certain extended ways? How can the technology be designed to mitigate the dangers? What kind of regulations need (or ought) to be put in place to control the use of the information?

D. Summing Up

Ethics has a role to play early in the process of technological development. Stahl and Rogerson (in [20]), the current and immediate past directors of the Centre for Computing and Social responsibility at DeMontfort University, acknowledge the difficulty of forecasting future developments, but through an examination of European Union projects isolate a number of ICT trends. These are; increasing computational power and decreasing size, new types of interfaces, network connection anywhere and anytime, and increased development of virtual places. These trends suggest ethical issues that require addressing, or at least, serious consideration. Goujon and Flick (also in [20]) are critical of the sharp divide that appears to exist between ethics and technology, and argue for ethical reflexivity, that is, the constant reassessment of ethical norms in response to the technology to economics, to politics and to gaps in regulation. The context of development is particularly

 ISBN: 978-1-61208-228-8

important and the non-neutrality of ICT requires emphasis. Inadequacies in both current risk assessment and technology assessment are highlighted.

IV. The case for eco-informatics

So, what lessons can eco-informatics take from this discussion? Think of eco-informatics as part of the larger environmental ethics movement. Think of eco-informatics as part of the larger information ethics movement. But, as discussed earlier, if eco-informatics is to stand as a meaningful term then we must search the eco-informatics literature for an example that can be used to highlight where ethical deliberations might be put.

I will use the paper, *What is meant by "decision-making" in the context of eco-informatics?* by János Fülöp, David Roth, and Charles Schweik [5] to show how an eco-informatics researcher might include ethics in their deliberations.

A. Decision Styles

The paper starts with the standard descriptions of decision making styles: Rational, Contingency Perspectives, Incremental, and Garbage Can Decision-Making. It is in the Rational style that ethics makes its only appearance in the whole paper. In the context of choosing the cheapest copier and vendor in which all the copiers are identical, they quote Rainey [15] saying that "to choose another vendor 'would invite others to question the [manager's] competence, ethics, or sanity"'. They suggest that it would be unethical to not take the cheapest offer.

However, it is not nearly so straightforward. There are some decision makers who consider themselves to be perfectly rational but would make their choice in the following way.

1) Determine the integrity of the Vendor. If it is acceptable,
2) determine the character of the Vendor's Representative. If the Representatives are acceptable,
3) determine the way in which the copiers are identical. Are they merely functionally identical? They cannot be absolutely identical without violating someone's copy or patent rights. Are the differences ethically relevant? If so which is the most relevant?
4) Having done the previous steps which copier makes the greatest contribution to society.

This decision making process is rational but, simply choosing the cheapest alternative is not necessarily the right choice. In fact, using this process it would be reasonable to 'question the [manager's] competence, ethics, or sanity' if they did not take ethics into account.

Similar comments could be made for each of the other styles; Contingency Perspectives, Incremental, and Garbage Can.

With Contingency Perspectives the emphasis is on 'judgment and intuition'; the decisions makers engage in'bargaining and political manoeuvering in their decision-making process'. Here, the authors could easily include ethical judgment and moral intuition [1] in these descriptions.

Incremental decision-makers 'choose to make less controversial, intermediary decisions to ensure some degree of success of achieving vague goals'. Vagueness, can be seen to be either a virtue or a vice. Some success can be seen as providing some benefit while minimising (usually) potential harm. Virtue and harm are both standard ethical concepts. This decision making style is already (if implicitly) framed in ethical terms.

Finally, the Garbage Can style is epitomised by 'waiting for an opportunity to advocate actions already planned'. Opportunism is usually seen in an unfavourable light. It is frequently considered unethical to be opportunistic as this has hints of using people, treating them as means to an end. Of course this is not necessarily true but the mere mention of this style is an opening to bring in the ethical notions of respect for persons, responsible decision making, and so on.

The authors end this section of their paper with:

> it is probably safe to say that most developers of eco-informatics tools or information hope that their work will be utilized in some form of rational decision-making processes or that at the very least, their tools and information are used to help inform incremental decision-making processes

Again, this a point at which the ethically sensitive eco-informatics researcher might add that 'their tools and information [could] be used to help inform' projects with ethical merit.

B. Optimization approaches

Here the authors attribute to Nemhauser the idea that optimisation 'can be applied in decision settings where there is a single criterion to base a decision on (such as cost)'(Nemhauser, et al. [13]). It would be relatively easy to include other criterion such as ethical worth, contribution to society, and so on.

Also, in this section, they refer to the development of 'GIS-based models'– a natural for eco-informatic systems. Given the recent problems that Google earth has had with privacy, identity theft, and the like it would be prudent to include such considerations in the constructions of such systems. See the Professional Ethics in engineering Forum [4] and Stephen and Cizek [16] for contemporary discussions of such ethical matters.

In the last example for this section of their paper, the authors are right to say:

> Another important point related to the various decision-making tools and approaches cited above as well as others falling in the domain of eco-informatics is that often variables used in computer-based models (such as multi-attribute decision models) are set to subjective values. These models may contain uncertainties, either because subjective scoring or are based upon some data or model output that contain some level of uncertainty. It is therefore an important question how the final ranking of alternatives are sensitive to the changes of input parameters contained within the decision model and how uncertainties are communicated to the user through analytic tools.

 ISBN: 978-1-61208-228-8

This is a perfect opportunity to introduce ethics into the discussion. It is not enough to simply say that communication with users is an important question. Here, they might say that it is not only an important question but rather it is a moral imperative to communicate the user so they can make the most fully informed choice, so they can give real, meaningful, informed consent.

C. Politics

The authors then touch lightly on the important topic of the use of eco-informatics for political ends.

> some scholars lament the fact that various eco-informatics-based tools, computer-based models and information are utilized as 'weapons in political and policy warfare,'while others accept this supporting role

There is insufficient space in this paper for a full examination of the ethical implications of this statement. Suffice it to say that introducing the ethics of such use is very important. Eco-informatics researchers ought to have strong well thought out moral stances to these kinds of uses of their work and these stances ought to be communicated through their writing.

D. Complexity

Also in this part of their paper the authors say that:

> Others worry about the complexity of eco-informatics-based computer models and decision makers' abilities to understand them.

This fits in nicely with the previous discussion that eco-researchers ought to strive to enable users of their work to be able to take informed decisions, to give informed consent. They could easily stress the ethics of such enabling writing.

E. Context

In the graph of the main part of their paper, the authors make the important point that:

> it is probably the case that developers of eco-informatics tools develop these methods without too much concern for the context in which such tools will be applied to decision making. Developers may focus more on getting the tool right (scientifically) and concentrate less on how the tool might be used or abused in particular decision-making settings

They present this in a neutral way: it ought not to go uncommented. They could to make the much stronger point that the developers have significant and important duties to the users, and society at large, for the responsible use of their efforts. The authors have an excellent opportunity to show the ethical urgency of changing this sort of attitude. Developers ought to be concerned about the context in which their tools are used.

V. Conclusion

I have shown that ethics does not get a serious examination by eco-informatics researchers and hence by social eco-informatics researchers. I have argued that ethical deliberation ought to be in the forefront of all researchers thoughts. Also, I have shown that with some reflection it is straightforward to put ethical deliberation into an eco-informatics research paper.

References

[1] R. Audi, *The Good in the Right - A Theory of Intuition and Intrinsic Value*, Princeton University Press, 2004.

[2] D. Collingridge, *The Social Control of Technology*, St. Martins Press, New York, 1980

[3] M. Croy, "Collingridge and the control of educational computer technology", *Society for philosophy and technology*, 1(3-4), 1996, available at http://scholar.lib.vt.edu/ejournals/SPT/, [retrieved: June 2012]

[4] Eng-Tips forum, *Google ,Maps/ Earth ethics question, thread765-215311*, Available from: http://www.eng-tips.com/viewthread.cfm?qid=215311, [retrieved: June 2012]

[5] J. Fülöp, D. Roth, and C. Schweik,*What is meant by "decision-making"in the context of eco-informatics?*, Available from: http://academic.evergreen.edu/projects/bdei/documents/decisionmakingsummary.pdf, [retrieved: June 2012]

[6] D. Gotterbarn, "A practical mechanism for ethical risk assessment a SoDIS inspection"in Herman T. Tavani and Kenneth Einar Himma (eds), *The Handbook of Information and Computer Ethics*, Wiley, Hoboken, NJ, pp. 429-471, 2008

[7] D. Johnson, *Computer Ethics*, Upper Saddle River, NJ. Prentice-Hall, 2001

[8] J.M. Keynes, *The Scope and Method of Political Economy*, London, Macmillan and co., 1890

[9] J.J. Kineman and K.A. Kumar, "Information as communication", *Proceedings of the 50th Annual Meeting of the International Society for the System Sciences*, http://journals.isss.org/index.php/proceedings50th/article/view/367/136, 2006, [retrieved: June 2012]

[10] A.V. Koshkariov, V.A. Tikunov, and A.M. Trofimov, '"The current state and the main trends in the development of geographical information systems in the U.S.S.R.", *International journal of geographical information systems*, Volume 3, Issue 3, 1989

[11] R. Lucas, What is there to be ethical about with Enterprising Computing?, *EC@UC* 15^{th} - 16^{th}, May 2012, Canberra, 2012

[12] J. Moor, "What is computer ethics?", *Metaphilosophy*, 16, pp. 266-275, 1985

[13] G.L. Nemhauser, A.H.G. Rinnoy Kan, and M.J. Todd, *Handbooks in Operations Research and Management Science: Volume 1 Optimization*, Amsterdam, North-Holland, 1989

[14] P. Nicholls, *Cross-Disciplinary Research and Education Post Copenhagen 2009*, The Royal Golden Jubile International Seminar Series LXX, p. 268, Graduate School of Business, Assumption University, 2009

[15] H. Rainey, *Understanding and Managing Public Organizations*, San Francisco: Jossey-Bass, 2003

[16] S.R.J. Sheppard and P. Cizek, "The ethics of Google Earth: Crossing thresholds from spatial data to landscape visualisation", *Journal of Environmental Management*, Volume 90, Issue 6, May 2009, pp. 2102-2117, ISSN 0301-4797, 10.1016/j.jenvman.2007.09.012. (http://www.sciencedirect.com/science/article/pii/S0301479708001151)

[17] C. Thornley, Stuart Ferguson, John Weckert, and F. Gibb, "Do RFIDs (radio frequency identifier devices) provide new ethical dilemmas for librarians and information professionals?", *International Journal of Information Management*, 31, pp. 546-555, 2011

[18] F. Villa, "New computer architectures as tools for ecological thought", *Trends in Ecology & Evolution*, Volume 7, Issue 6, pp. 179-183, 1992

[19] J. Weckert, "An Approach to Nanoethics", in Graeme Hodge, Diana Bowman and Karinne Ludlow (eds), *New Global Regulatory Frontiers: The Age of Nanotechnology*, Edward Elgar, 2007

[20] J. Weckert and R. Lucas *Professionalism in the Information and Communication Technology Industry*, Canberra, ANU ePress, *forthcoming* 2012

[21] M. Williams, *Nano-based RFID tag, you're it*, Available from: http://www.media.rice.edu/media/NewsBot.asp?MODE=VIEW&ID=13899, [retrieved: June 2012]

[22] N. Wiener, "Some moral and technical consequences of automation", *Science*, 131, pp. 1355-1358, 1960

 ISBN: 978-1-61208-228-8

Counterinsurgency Through Civil Infrastructure Networks

David Walker, Simon Reay Atkinson
Complex Civil Systems Research Group
Faculty of Engineering and IT
University of Sydney, NSW 2006, Australia
[david.walker, simon.reayatkinson]@sydney.edu.au

Liaquat Hossain
Department of Informatics
Lund University, Sweden
Liaquat.hossain@ics.lu.se

***Abstract*—We investigate opportunities and risks to counter-insurgency (sometimes referred to as COIN) that are inherent in the civil networks surrounding infrastructure projects. It is argued here that a) successful counter-insurgency largely comprises the 'engineering' of robust and trusted civil networks that are capable of re-channelling insurgent designs; and b) introduction of externalized-exclusive control networks poses a significant risk to such network development. In this conceptual paper, we propose a number of network models, each hypothesizing a risk and / or opportunity. These models will be tested and refined using a case study methodology that draws on documentary evidence and interviews with subject matter experts. We propose two fundamental relationships: Coordination by rule and control (CRC), and; Collaboration by social influence (CSI). CRC is based on mechanical (rule and time based) structures, while CSI is based on organic (informal, trusted and shared aware) social networks. We posit that COIN is primarily a CSI interaction / exchange that can be supported (but also obstructed) by CRC. We therefore propose that the design of successful COIN operations must adhere to the hybrid notion providing an interface between CRC and CSI.**

Keywords - civil networks; infrastructure; insurgency; coordination; collaboration

I. BACKGROUND AND CONTEXT

A. Introduction

Organisational communication literature maintains that hierarchical structures, e.g. organograms, provide a superficial representation of how work actually gets done [1]. Here, we posit (after Hossain and Wigand [2]) that organisations need to be seen as dynamic (elastic and plastic) social-influence networks. In these collaborative [3] networks of complex operations, requiring tacit knowledge exchange [4], which is achieved through social (and in this regard civil) interactions beneath and sometimes masked by the formal hierarchical organisational chart. Organisational co-adaptive [5] viability in maintaining operational effectiveness and efficiency may therefore be largely dependent on how we synergistically socialise and capitalise 'our' formal (hierarchical) and informal (social) networks to achieve shared common goals. In this paper, we set the context as relating to counterinsurgency which we then explore through the mediation of hierarchical-formal and informal-social networks as applied to civil infrastructure projects. In this paper we also ask whether the success and/or failure related to COIN is attributed to the conceptual misunderstandings of how we provide synergy between hierarchical formal-control-rule (CRC) and informal-collaborative-social-influence (CSI) networks. In this regard, we distinguish between Command, as in Leadership and Control, as in Management and consider, after Reay-Atkinson and Moffat, that: Control is a function of rules, fidelity, time and bandwidth whereas command is a function of trusts, shared awareness, influence and agility [6].

Securitization emerged along with Contractorization from the 'politics' of Privatization put in play in many democracies from the 1980s onward. Whereas contractorization (of security) was manifested through Private Finance Initiatives (PFIs) and Private Security / Military Companies (PS/MCs); privatization was more often introduced through Public-Private Partnerships (PPPs), for example the privatization of British Defence Research into Dstl and QinetiQ. Building on Grotius [7], we posit three interconnected and interactive conflict phases: jus bello (justice to war); jus bellum (just war) and jus post bellum (justice after war). In our understanding, the trusts established before during and after conflict underpin the public assurances, trusts, safety and security subsequently established. No one phase trumps the other. This has specific implications for how counter-insurgencies are engaged, or not. The Civil Infrastructure Networks we identify in this paper are not limited in scale and connect the social (people) with civil programmes; including IT, transport, Cyber-, water and energy networks. Post Privatized forces tended to optimise out these very skills / trusts through PPPs, PFIs and PS/MCs. This had a number of effects. The first was to create warrior cast structures with front line troops at the top of the pyramid and engineering / logistics even health care at the bottom. Secondly, it created incoherence in command / collaboration between divided force structures for which many of the 'enablers' were now under contract, e.g. tank transportation. Thirdly it created a reliance of the fighting force on structures that a) could not necessarily be put 'under disciplined control' and / or b) expected to be collaboratively deployed into hazardous areas. Fourthly, it created private and sometimes competing armies – very often employed by other government departments (OGDs), such as the UK Foreign and Commonwealth Office. Fifthly, the opportunity to contract one's own security / reconstruction programmes removed the need for healthy local collaboration and mutuality between government departments, e.g. between the state department and the Department of Defence.

This paper is organised as follows: first, we discuss COIN and its context; secondly, we discuss the application

 ISBN: 978-1-61208-228-8

of social networks in modelling risks and opportunities for COIN; thirdly, we discuss the social influence model within the context of COIN and provide conclusions, implications and future directions.

B. Counterinsurgency

US Military doctrine considers an insurgency to be: 'an organized, protracted politico-military struggle designed to weaken the control and legitimacy of an established government, occupying power, or other political authority while increasing insurgent control' [8]. Counterinsurgency is every effort to stop an insurgency, once started. The primary objective of COIN is to maximise the support of the civilian population for the legally constituted government [8-13]. Insurgents depend on the civilian population for sustenance, shelter, concealment and recruitment. Insurgents (almost always clandestine in nature) employ a number of strategies to shift the objective demographic in their favour, including political and judicial subversion, terrorism, guerrilla warfare, organised crime / the Black economy (extortion, corruption, smuggling etc), propaganda, and public service provision (including governance, security and judiciary) [8, 13]. Technological advances and demographic trends over the last several decades have significantly changed the nature of conflict. In a contest between a small-scale, agile, poorly funded but technically literate and socially aware clandestine movement and a large technologically advanced and hierarchical rule-based military, these trends have tended to support the former. They include urbanisation, globalisation, proliferation of NGOs and other civilian organisations, ubiquity of news media, revolutionary advances in information technology, social media, increased lethality of highly mobile weaponry, and global [religiously based] insurgency [14].

Contemporary insurgencies (considered by this paper) may lack the ideological appeal of classical insurgencies against 'local' colonial forces and therefore rely on causing existential intervention by government / coalition forces to generate localised conflict – essentially localising the conflict for 'home' advantage. Kilcullen [10] develops the idea of an 'accidental guerrilla syndrome'. He theorizes that 'the accidental guerrilla emerges from a cyclical process that takes place in four stages: infection, contagion, intervention, and rejection.' Importantly, this insurgent strategy assumes counterinsurgents will generate interstitial conflict when they enter local networks.

Effective COIN requires coherent and collective (unifying) input from a multiple stakeholders including military, other government agencies / departments, non-government organisations, private contractors, and most importantly any functioning civil networks remaining embedded in local populations. Yet usually there is no remaining control structure and legal authority by which networks can be coordinated – even and although the local population continue to feed and fend for themselves. *The inescapable corollary is that coordination of counterinsurgent organisations and local civil networks needs to be achieved by something other than rule and control. We call this collaboration by social influence (CSI).* This paper addresses two questions: 1) How can the efforts of such a diverse array of actors be influenced and / or coordinated? 2) What are the implications for the internal networks of counterinsurgent organisations? To address these questions, a number of network models are developed, each representing a key opportunity or risk to COIN.

Civil infrastructure projects have a number of crucial characteristics that make them ideal for COIN and ideal for research into the dynamics of civil networks. Firstly, they provide justification for large scale information transfer to and from the population. Secondly, project stakeholders and their networks cover almost every facet of society and outside influence; including military forces, end-users, contractors, suppliers, NGOs, local government, tribal and religious leaders, political opposition, financial service providers, donors, and land owners. Thirdly, projects have clear boundaries and objectives. Fourthly, they provide economic stimulus for development and maintenance; and fifthly, the tangible benefits of the infrastructure itself remains. Most importantly from a research perspective, infrastructure projects: a) give rise to a large quantity of reliable network data; and b) provide identifiable boundaries.

II. Modelling Opportunity and Risks to Counterinsurgency

A. Social Networks and Their Impact on Risk and Opportunity

Opportunities for CSI may be considered a type of social capital (SC). Coleman defines SC as 'a variety of different entities, with two elements in common: they all consist of some aspect of social structures, and they facilitate certain actions of actors' [15]. Coleman provides a few clarifications that are particularly important to the present study: 1) SC can constrain as well as enable action; 2) SC that is beneficial to one activity may be detrimental to another; 3) SC, like all capital, need not be utilised; and 4) SC created for one purpose can be used for another. With regard to stabilising reform in COIN, it may be hypothesised that the 'low hanging fruit' is made up of two basic types of existing social capital: 1) that which could have a stabilising effect but is currently not being utilised or is being utilised for some other purpose; and 2) pernicious structures that constitute relatively trivial social capital and will therefore not be vigorously defended. This 'low hanging fruit' represents the opportunities for reform that counterinsurgents find so difficult to identify. Social network analysis is one way of modelling social capital.

Social network analysis (SNA) is 'a distinct research perspective', which includes 'theories, models and applications that are expressed in terms of relational concepts and processes' [16]. A network is made up of actors (individuals, organisations, or some other social unit) and relational ties between them. SNA attempts to model and predict patterns of relational ties (network structures), and to understand the causes and effects of these patterns. Laumann et al [17] observe two ontological perspectives of networks: realist and nominalist. From a realist perspective, a 'network is treated as a social fact only in that it is consciously experienced as such by the actors composing it'. With a nominalist approach

 ISBN: 978-1-61208-228-8

the analyst 'imposes a conceptual framework constructed to serve his or her own analytic purposes' [17]. This paper adopts the nominalist approach, which provides greater flexibility in network definition and enables the analysis of disconnected networks.

1) Network Dynamics. Balance Theory developed by Austrian psychologist Fritz Heider specifies that a triad (group of three actors) is balanced if all ties are positive or if two are negative and one is positive [18]. In SNA, the well observed tendency for triads to become 'balanced' is called transitivity [16]. In the present study we are particularly concerned with the effect of new actors entering a social network. If actors A and B are friends and new actor C becomes an enemy of A, by the theory of transitivity it can be expected that B and C will become enemies. This is a risk for C. Less obviously there is also risk to C in creating positive (friendship) ties. If A and B are enemies and new actor C becomes friends with A, it can be expected that B and C will become enemies. Homophily, which is the tendency of people to make connections with others that are similar to themselves [19, 20], may amplify or dampen the effects of transitivity.

The dynamics of conflict pacification and escalation within networks are crucial. Humans tend towards a 'tit-for-tat' strategy in response to acts they consider unreasonable [21]. In relation to the actions of others, humans have a tendency to perceive their own contributions more significant, their own gains more deserved, and their own losses more unjust. The logical (and observable) outcome of these human conditions is a positive feedback cycle within which animosity and conflict (once initiated) escalates between two actors. Importantly, an actor responds to his/her *perceptions* of another actor's behaviour and intent, so the positive feedback loop can begin without any *actual* acts (or even intent) of aggression. Transitivity also dictates that conflict relationships may cause other actors in the network to become conflicted.

Trust, like conflict, may propagate or collapse exponentially through a network. Dasgupta [22] concludes that "trust is a public good, a social lubricant which makes possible production and exchange...[It] is based on reputation...acquired through behaviour over time in well-understood circumstances." Trust is dependent on social networks, and it "is this interconnectedness which makes trust such a fragile commodity." Individuals place trust according to their (often intuitive) calculations of risk and return. A significant element of that risk exists in relation to the reliability of their information on a potential exchange partner. Therefore, significant gains in social capital are made by the provision of reliable information to all parties. This may lower the cost of trusting below some point of criticality and so generate a chain reaction (or cascade) of trusting relationships. Such an effect is a worthy objective of temporarily intervening actors.

Knowing how people 'connect' makes collaboration work better; see Mintzberg et al [3]. They suggest that collaboration may ultimately depend on trust. Collaboration depends on the ability to trust each other, and to appreciate one another's expertise. Perhaps, surprisingly, they argue, the best collaboration may be the least realised as collaborative, giving the example of interdepartmental collaboration for new product development. In the best of such collaborations of joint learning whilst designing, people focus intently on 'shaping' but may not even realise they are collaborating, so that shifting their focus to formal techniques of collaboration may, in fact, reduce their capacity and propensity to collaborate. Building on this concept of trust, Marsh [23] suggested the following definition of trust:

> '…trust, (or symmetrically, distrust) is a particular level of the "subjective probability" with which an "agent" will perform a "particular action", both before he can monitor such action (or independently of his capacity to monitor it) and in a "context" in which it "affects" his own action.'

2) Information and Control. Building on the theory of transitivity, Granovetter [24] proposed that most novel information is attained through weak ties. The Strength of Weak Ties (SWT) theory is based on the relatively simple logic that a person's strong ties most likely lead to people that are also strongly tied (due to transitivity). This means that information coming to an individual through one strong tie is likely to be the same as information coming through another strong tie. Weak ties on the other hand are more likely to lead to people that would otherwise be only distantly connected or not connected at all. In Granovetter's seminal study he provides empirical evidence for this theory by showing that people are more likely to find a job through a weak tie than a strong one.

Ronald Burt's structural holes (SH) theory also seeks to explain how network structure contributes to an individual's access to novel information. Structural holes are the gaps in a social network between two actors that are not connected. If another actor manoeuvres into that gap by forming a relationship with each of those actors, he or she is then in a position to control the flow of information and resources between them [25]. One's network is *effective* to the extent that it reaches many other actors (through primary contacts and their close ties) and *efficient* to the extent that the ratio of total contacts reached to total primary contacts is large. The *effective size* of one's network is deemed to be the total number of non-redundant contacts. Non-redundant contacts are those between which there is a structural hole and therefore those that provide access to novel information.

Burt considers the SH explanation superior to that of SWT because "the causal agent in the phenomena is not the weakness of a tie but the structural hole it spans…[and] the weak tie argument obscures the control benefits of structural holes" [25]. The control benefits of structural holes are indeed very important to SH theory and very important to this study. If actor A spans a hole between actors B and C, actor A has control benefits to the extent that B and C make mutually exclusive demands (or requests) on A. The traditional auction is a simple case in which an offer from B is used to raise the offer from C. For most negotiations, however, it is

 ISBN: 978-1-61208-228-8

secondary holes that infer most control. Secondary holes exist between an actor's contacts and others that could replace that contact. Where one actor could substitute for another, they are considered to be *structurally equivalent*. The simplest example is a market. A buyer has power to drive the price down to the extent that there are multiple sellers competing for that sale. The competing sellers are structurally equivalent. Structurally equivalent actors can avoid being played against one another by coordinating their actions. This type of coordination takes many forms within the economy, e.g. labour unions and price fixing cartels.

Network exchange theory provides an alternative way to examine power relations and brings to light at least one counterintuitive insight that is not easily identified through SH theory.

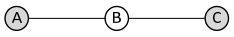

Figure 1. Powerful Central Actor

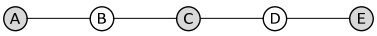

Figure 2. Weak Central Actor

While "centrality measures have typically been used as indicators of power, influence, popularity and prestige" [26, 27], Markovsky et al [28] demonstrate that under certain conditions central actors may be considerably weaker than those that are not particularly central. In Figure 1 actor B is the most central and the most powerful in negotiations. In Figure 2, however, the most central actor (C) is less powerful than B and D. This is because B and D can extract very high returns from A and E respectively and so reduce their dependence on C, possibly cutting C out altogether: social power can thereby depend on connections to others in weaker positions.

Network models of power are highly relevant to our understanding of COIN. As well as competition with the insurgency, counterinsurgents can find themselves in competition with one another (they are structurally equivalent and not coordinated). 'Competition' may also come from other global powers posturing for influence in a region and from NGOs reluctant to cooperate with any central or Government or International authority, or even from local coalition allies. In Afghanistan, there were examples of some Coalition partners refusing / being unable within their [control] rules-of-engagement to collaborate with local partners due to their previous criminal / conflict linkages and records. In some cases, local partners turned to another Coalition ally who, while maintaining their political and bureaucratic influence, also inadvertently preserved the power base and rationale for the insurgency in the first instance.

B. *Network Models of Social Influence*

Building on the context and theory, in this section we present a number of network models that hypothesise on the dynamics of social influence during COIN. Through these models we identify risks and opportunities inherent in the structures of social networks and the actions of key players. In accordance with the nominalist approach, networks may be defined on any conceivable tie. Conceiving possible networks is a crucial task of domain experts. Productive use of these models requires close cooperation between domain experts (practitioners) and theoretical experts (academics). The inevitable shortcomings of each type of expert acting without the other has been recognised as an important barrier to research [29].

1) Unintended Exclusion (the danger of strong ties). Figure 3 presents network dynamics in response to counterinsurgents entering a network and forming a strong tie with one actor while inadvertently (or deliberately) excluding structurally equivalent others.

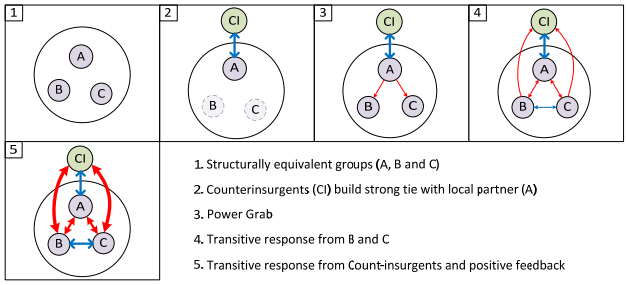

Figure 3. Unintended Exclusion

From a 'realist' perspective Figure 3 may show three tribes or three warlords (A, B and C). However this is only one of many possibilities, even if a valid interpretation. From a nominalist perspective it is posited that such a dynamic might be observed within any network built on inferred power relations.

Structural equivalence provides the capacity and intent to perform some function and exercise some control. Networks that make up structurally equivalent actors vary depending on the type of tie under consideration. Networks may be formed by ties that, for example, determine land tenure, opium production, trucking (or other goods and services), spiritual guidance, political leadership, judicial authority and policing. Each type of tie produces a different network signature which may be observed. We suggest that insurgents may be 'spread' throughout local networks. Insurgents seek to infiltrate and suborn ('infect') local networks by breaking down their immune systems and occupying traditional roles, marrying into tribes, creating business partnerships and converting susceptible minds to their cause.

From this model it is hypothesised that: 1) strong ties increase the likelihood of negative response from unknown others; 2) Weaker ties with all actors may result in more positive (or neutral) ties (collaboration) with unknown others; but 3) Maintaining weaker ties requires counterinsurgents to look beyond the rather more obvious / recognisable (like) and accessible / attractive strong ties generally available.

2) Self Fulfilling Prophecy. The network model presented in Figure 4 represents the inadvertent escalation

of conflict that may begin with nothing more than a perceived display of hostility. Instigating this dynamic through deception and intimidation is a key strategy of insurgents. The presence of counterinsurgents provides the insurgency with 'initiation' opportunities. Through their detailed understanding of collaborative-co-adaptive and cooperative-competitive relationships, insurgents can shape the ecology to their advantage by convincing otherwise peaceable citizens that their way of life is under attack. The easy thing for counterinsurgents to do is to target an insurgent. Yet often an insurgent may also be part of other Black Economy type activities essential for supporting the well-being of the local population. In this instance the counterinsurgents will frequently be better off observing and not contesting these agents; while working to convince networks to reject more malign or un-reconcilable actors.

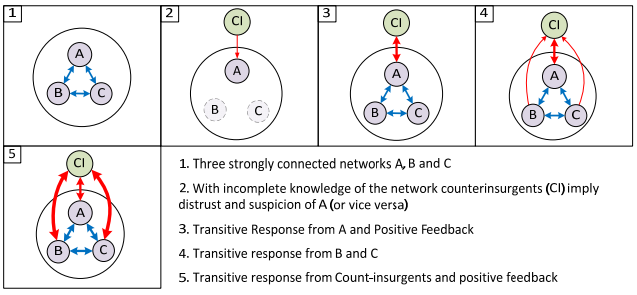

Figure 4. Self Fulfilling Prophecy

3) Coercion. A key task of COIN is to identify local civil system-networks that are capable of resisting / being immunised against coercion. In this model (Figure 5), coercion is perceived as a star network with a coordinated coercer at the centre connected to a number of other networks that might (if sufficiently collaborative) be able to reject (or rechannel the designs of) the coercer.

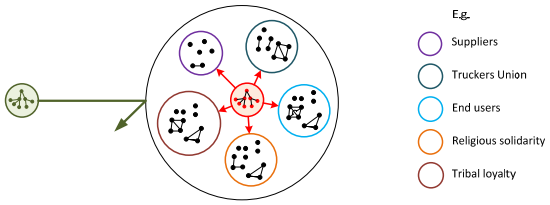

Figure 5. Coercion

Each of the networks is defined on a different tie, so a single actor may exist in more than one network. We assume that within any given population there exists a number of networks through which power and influence is distributed. Networks will reject malign actors to the extent that they are capable, motivated and immunised to do so. The capability of each network to reject (or rechanneled the designs of) malign actors is a function of the network's coordination, social power (implicit in the type of tie that forms the network), and reach (collaboration). We assume motivation is based on cost / benefit appreciation with some appropriate weighting for risk aversion / acceptance.

The tacit knowledge necessary to build this model is a *critical information requirement* of COIN. Importantly, this is information about licit networks, not clandestine insurgent networks. We assume coercion exists. The information we seek to attain and distribute is that which will influence existing networks to collaboratively coordinate and exclude malign elements. Military organisations (particularly intelligence assets) have deeply entrenched tendencies towards information control; often precluding the kind of knowledge exchange and shared awareness required to build such a model. See Flynn et al [30] for a cathartic account of military intelligence failures.

The potential of a given network to improve its capability, capacity and intent is also critical. Some networks are fixed, while others may have potential to dynamically change size, power-relationships and purpose. An example may be the National Solidarity Programme (NSP) in Afghanistan. It was designed to decentralise the control of civil infrastructure projects and disaggregate legitimacy and control to the local level. In the absence of trusted [nationwide] judicial and political systems, it became a common source of impartial judgment so a) preventing and b) resolving local disputes through collaborative social influence [31].

4) Ideal Instance of CSI. Figure 6 represents a network dynamic that is a key objective of COIN. Based on theory and the risk profiles previously identified, we posit that opportunity for these network dynamics may emerge from a number of conditions. Firstly, counterinsurgents are sufficiently coordinated to avoid being played against one another - so enabling them to make explicit and credible promises of withdrawal. Secondly, counterinsurgents avoid empowering any single local network that is structurally equivalent to multiple others. Thirdly, local civil networks are identified which: a) have a significant collaborative interest in sustained support from counterinsurgents, and b) are capable of rejecting / being immunised against (or rechanneling the designs of) malign actors. Fourthly, hostile action or intent (and the impression thereof) towards any individual or group is minimised through shared awareness and the threat being seen as an attack on the whole.

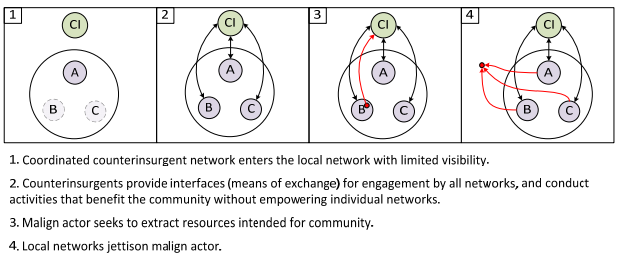

Figure 6. Local Networks Rejecting Malign Actor

Infrastructure projects provide the ideal context to study and exercise this process. Black Market / criminal elements

will seek to extract resources as long as they can portray themselves as enemies of counterinsurgents as opposed to the public good. Such a 'posture' requires a very high standard of project management (to identify, prevent and prosecute hostile activity) while collaborating effectively with the civil networks within which some black marketeers / criminals will inevitably seek to hide. This scenario is an example of CRC (to produce evidence with high legitimacy) supporting CSI (to influence perceptions and motivate actions by local networks).

III. CONCLUSION AND FUTURE RESEARCH DIRECTION

Drawing on theories of social capital, strength of weak ties, structural holes, network balance, and network exchange we have developed a set of models of network dynamics in COIN. The hypotheses of this paper are implicit in these models and the deductions drawn from them. We recognise the underlying moral requirements underpinning the effective deployment and use of Armed Forces in a counter-insurgency. And this moral / ethical underpinning – identified in just war theory – underpins not only the success of the operations but also the ability of our Armed Forces and the local population to recover from instability. In this respect we can see the significant reduction in the extent and subsequent impact of PTSD on deployed forces. We see civil infrastructure networks – at all scales from the Cyber- to 'bridge building' – as underpinning a successful counter-insurgency and re-connecting shattered communities after conflict; an example being the Mostar Bridge.

From this conceptual starting point, the study will progress through two phases. An interview protocol will be developed and an initial set of interviews with subject matter experts will be used to test that protocol and further refine the theoretical models. The second phase will be a rigorous testing of these models through case studies selected for theoretical replication; utilising the case study framework developed by Yin [32]. This qualitative approach is considered most appropriate because it allows us to "retain the holistic and meaningful characteristics of real-life", which is important when investigating complex social phenomena [32].

REFERENCES

[1] R. Stacey, "Creative Organizations: The Relevance of Chaos and Psychodynamic Systems," Social creativity, vol. 2, 1999, pp. 61.

[2] L. Hossain and R. T. Wigand, "Ict Enabled Virtual Collaboration through Trust," Journal of Computer-Mediated Communication, vol. 10(1), 2004.

[3] H. Mintzberg, J. Jorgensen, D. Dougherty, and F. Westley, "Some Surprising Things About Collaboration-Knowing How People Connect Makes It Work Better," Organizational Dynamics, vol. 25(1), 1996, pp. 60-71.

[4] S. Reay-Atkinson, S. Lesher, and D. Shoupe, Information Capture and Knowledge Exchange: The Gathering, Testing and Assessment of Information and Knowledge through Exploration and Exploitation, 2009, DTIC Document.

[5] S. Reay-Atkinson, A. Goodger, N. Caldwell, and L. Hossain, "How Lean the Machine: How Agile the Mind?," The Learning Organization, vol. 19(3), 2012, pp. 183 - 206.

[6] S. Reay Atkinson and J. Moffat, "The Agile Organization," Publication of the US Department of Defense Command and Control Research Program (CCRP), 2005.

[7] H. Grotius, De Jure Belli Ac Pacis - on the Law of War and Peace. 2nd ed. Paris (2nd ed. Amsterdam 1631), 1625.

[8] US-Army, Field Manual 3-24 Counterinsurgency Washington, DC: HQ Department of Army, 2006.

[9] D. Kilcullen, "Counter-Insurgency Redux," Survival, vol. 48(4), 2006, pp. 111-130.

[10] D. Kilcullen, The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One. Melbourne: Scribe Publications, 2009.

[11] B. Hoffman, "Insurgency and Counterinsurgency in Iraq," Studies in Conflict & Terrorism, vol. 29(2), 2006, pp. 103-121.

[12] K. I. Sepp, Best Practices in Counterinsurgency, 2005, DTIC Document.

[13] D. Galula, Counterinsurgency Warfare: Theory and Practice. Westoprt, CT: Praeger Security International, 1964.

[14] ADF, Complex Warfighting: Future Land Operating Concept. Canberra 2006.

[15] J. Coleman, "Social Capital in the Creation of Human Capital," American Journal of Sociology, vol. 94(S1), 1988, pp. 95.

[16] S. Wasserman and K. Faust, Social Network Analysis: Methods and Applications. Cambridge University Press, 1994.

[17] E. O. Laumann, P. V. Marsden, and D. Prensky, "The Boundary Specification Problem in Network Analysis," Research methods in social network analysis, vol. 61, 1989, pp. 87.

[18] F. Heider, "Attitudes and Cognitive Organization," The Journal of Psychology, vol. 21(1), 1946, pp. 107-112.

[19] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a Feather: Homophily in Social Networks," Annual review of sociology, 2001, pp. 415-444.

[20] P. F. Lazarsfeld and R. K. Merton, "Friendship as a Social Process: A Substantive and Methodological Analysis," Freedom and control in modern society, vol. 18, 1954, pp. 18-66.

[21] R. Axelrod and W. Hamilton, "The Evolution of Cooperation," Science, vol. 211(4489), 1981, pp. 1390-1396.

[22] P. Dasgupta, "Trust as a Commodity," Trust: Making and Breaking Cooperative Relations, electronic edition, Department of Sociology, University of Oxford, 2000, pp. 49-72.

[23] S. Marsh, "Formalising Trust as a Computational Concept," University of Stirling, Department of Mathematics and Computer Science, PhD dissertation, 1999.

[24] M. Granovetter, "The Strength of Weak Ties," American Journal of Sociology, vol. 78(6), 1973, pp. 1360-1380.

[25] R. S. Burt, Structural Holes : The Social Structure of Competition. Cambridge, Mass: Harvard University Press, 1992.

[26] J. Scott and P. Carrington, The Sage Handbook of Social Network Analysis. Sage Publications Ltd, 2011.

[27] L. Hossain, A. Wu, and K. K. S. Chung. Actor Centrality Correlates to Project Based Coordination. 2006. ACM.

[28] B. Markovsky, D. Willer, and T. Patton, "Power Relations in Exchange Networks," American Sociological Review, 1988, pp. 220-236.

[29] R. C. van der Hulst, "Terrorist Networks: The Threat of Connectivity," The SAGE Handbook of Social Network Analysis, 2011, pp. 256.

[30] M. Flynn, M. Pottinger, and P. Batchelor, "Fixing Intel: A Blueprint for Making Intelligence Relevant in Afghanistan," Center for a New American Security, Washington DC, Jan. 2010.

[31] Economist, "Democracy in Afghanistan - Wise Council," in The Economist, 25 Mar. 2010.

[32] R. K. Yin, Case Study Research: Design and Methods. Vol. 5. Sage publications, INC, 2009.

 ISBN: 978-1-61208-228-8

How Much do Digital Natives Disclose on the Internet – A Privacy Study

Erik Buchmann, Klemens Böhm
Karlsruhe Institute of Technology (KIT), Germany
{erik.buchmann|klemens.boehm}@kit.edu

***Abstract*—With the advent of social services on the Internet that encourage the disclosure of more and more personal information, it has become increasingly difficult to find out where and for which purpose personal data is collected and stored. The potential for misuse of such data will increase as well, e.g., due to the ongoing extension of social sites with new features that make it more appealing to reveal personal details. In order to research and develop approaches that give way to privacy on the Internet, it is important to know which kind of information can be found, who has been responsible for publishing it, the age of the information, etc. This paper describes a user study about the personal information available about Digital Natives, i.e., young people who have grown up with the Internet. In particular, we have guided 65 undergraduate students to search the web for personal information on themselves by using various search engines. Our students have completed 302 questionnaire sheets altogether. We have analyzed the questionnaires by means of statistical significance tests and cluster analysis. As a part of our results, we have found out that today's personal search engines like 123people.com do not find much more information than general-purpose search engines like google, and that today's Digital Natives are surprisingly aware of the information they are willing to disclose.**

Keywords–privacy; user study; digital natives.

I. INTRODUCTION

At this moment, more than 200 social networking sites exist [1] that encourage the disclosure of personal information about the daily life, hobbies and interests, work-related information, etc. Furthermore, more and more classical web forums, photo-sharing portals, news portals and other sites offer "social" features, e.g., allow to share and comment digital objects of interest. In consequence, it has become more and more tempting for individuals to disclose personal details, with unpredictable consequences for privacy.

The potential for misuse of such data depends on how it can be found. While general-purpose search engines like Google or Bing search for any kind of information, people-search engines like 123people or Yasni are tailored to search for personal details, e.g., from web forums, social network sites or commercial portals.

In this article, we investigate which personal information is typically available on the Internet. We study who has published this information and how our participants assess the impact of its availability on their privacy. From related research (see Section II) and privacy issues on public media, we have developed four classes of research questions:

A1: Information characteristics This class subsumes the extent of personal data available on the Internet, the age of the data and parties who have uploaded it.

A2: Search characteristics This class covers the influence of the search engine and the search terms on the search result, and the amount of ambiguous information.

A3: Impact on privacy This class considers if people are content with the fact that the personal information they have found on themselves has been uploaded and is available on the Internet, and how sensitive they deem this information.

A4: Patterns and rules The fourth class addresses relationships between the characteristics of the information. In particular, we look for clusters of people with similar privacy attitudes and privacy behavior.

We have conducted a user study with educated Digital Natives, i.e., with people grown up with the Internet. Digital Natives are relevant for our study, because those people have integrated the Internet into their daily life. A very large share of the adolescents and young adults belong to this group. We have decided in favor of a qualitative study, i.e., a very detailed questionnaire and an intensive supervision of the study participants. Over a period of three years we have guided 65 undergraduate students of computer science to search for personal information about themselves by using various search engines. We have asked them to state who has uploaded this information, the age of the data, who would be able to find it, etc.

As a part of our results, we have found out that today's Digital Natives are very aware of the information they are willing to disclose. Nevertheless, despite the fact that names can be ambiguous ("John Smith"), all of our participants found at least some information about themselves on the Internet, and they disagreed or strongly disagreed with the availability of about one fourth of this information.

Paper Structure: The next section reviews related work. Section III describes our study methodology. The study is presented in Section IV, followed by a discussion of the study results in Section V. Section VI concludes the paper.

An extended version of this paper that includes more analyses and discussions is available as a technical report [2].

 ISBN: 978-1-61208-228-8

II. Related Work

In this section, we explain the privacy paradox, we outline studies on Internet privacy in different use cases, and we discuss privacy perception and user categories.

Privacy Paradox: Our survey is motivated by the privacy paradox [3]: This paradox means that the attitude towards privacy and the daily behavior of individuals is inconsistent in many cases. For example, a study about anonymous and personalized gift cards [4] shows that people tend to assign a high price to the protection of a certain information, but in fact accept a much lower price to actually sell the same information. In contrast, we want to find out if there is a discrepancy between the personal information Digital Natives have explicitly published and the information they would tolerate to be disclosed. The privacy paradox can be modeled as a function of costs and benefits, which is maximized by each individual [5]. The costs include the risks of identity theft, marketing, stalking or negative reputation. Benefits include social aspects like relationships, collaborations, friendships or positive reputation in general. Related to the privacy paradox is the privacy awareness, i.e., the individual attention and motivation regarding the whereabouts of personal data. Privacy awareness influences individual decisions about publishing data [6].

Studies on Internet Privacy: Comparative privacy studies consider different use cases on the Internet:

Social Networks A study on information disclosure in social networks like Facebook or Myspace relates experience and behavior of users to the amount of private information that is disclosed [7]. Another study focuses on the privacy settings that control which information from the personal profile is shown to others [8].

eCommerce Privacy studies on customer data in eCommerce focus on the relationship between privacy and sales. A customer cannot observe if an online dealer follows the privacy policy on the shopping web site. Thus, a study [9] investigates the trust of the consumers in the willingness and ability of the dealer to handle personal data with care.

Personalization Many commercial web sites generate customer loyalty by personalization. This requires the customer to reveal personal details. The tradeoff between personalization and privacy is known as the online consumer's dilemma, which has been studied according to user value [10], transparency and willingness [5], and other impact factors [5].

The studies show that users tend to reveal information only if they see a direct use for it. For example, customers of a web shop do not disclose religious information [11]. This is important for our survey, because it shows that Internet users do not publish information indiscriminately.

Privacy Behavior and Privacy Perception: Surveys [12] about privacy behavior investigate the relationship between the perception of risks [13], e.g., identity theft, and the use of privacy-enhancing technologies. The studies show that the perception of privacy risks varies widely, but privacy behavior has been comparable among all participants.

Categories of Users: We are interested in identifying user groups that differ with respect to the personal information available on the Internet. An email survey of Internet users [14] has computed a score for privacy concerns on the Internet from questions about typical situations, e.g., if an individual registers for a company web site when receiving an unsolicited email about a new product. The survey has identified the categories "unconcerned user" (16%), "circumspect user" (38%), "wary user" (43%) and "alarmed user" (3%). Studies that directly inquire the privacy behavior from the users are prone to the privacy paradox. Our study in turn looks at this problem from a different perspective: we analyze personal information disclosed on the Internet.

III. Methodology

In this section, we compile concrete research questions and we describe our study methodology.

A. Research Questions

To investigate the personal information that is available on the Internet, we have come up with specific research questions, as follows:

A1: Information characteristics
- How much personal information is available?
- How old is the information?
- Who has made the information available?

A2: Search characteristics
- Which search terms have yielded most information?
- How much does the search result depend on the search engine?

A3: Impact on privacy
- Have our participants been surprised to find a particular piece of information?
- Had our participants given permission to upload the information?
- How sensitive do the participants deem the information they have found?
- Do the participants approve that this information is available on the Internet?
- Who is able to find which kind of information?

A4: Patterns and rules
- Do groups of individuals with different privacy perception and behavior exist?

B. Study Participants

We have tested our research questions on educated Digital Natives, i.e., on people who have grown up with the Internet, for two reasons. First, these individuals use the Internet frequently, and they are aware of the social benefits of sharing personal information, e.g., to keep contact with friends and relatives, or to find individuals with similar interests and

 ISBN: 978-1-61208-228-8

attitudes. Second, Digital Natives can be assumed to be able to develop strategies, e.g., using different pseudonyms and email addresses for different purposes, to prevent someone from learning personal details which are not for the eyes of others. We have conducted our study with 65 German undergraduate students of computer science. Since we had announced an anonymous study and demographic data is a quasi identifier [15], we did not collect such information.

C. Study Procedure

We have conducted our study in three tranches with different participants over a period of three years. In the first step of each tranche, we have described the purpose of the study to our participants. Furthermore, we have handed out a guideline how to search for personal details on the Internet by using different search engines, and by refining the search term if a search returns only results that do not have any relationship to the searcher.

In a second step, we have handed out a number of identical questionnaires to each participant. We have guided our participants to search for personal information, i.e., we have provided hints and support if necessary. We have asked our participants to answer one questionnaire sheet for each distinct search result, i.e., each answer sheet has been obtained using a different set of search terms and/or a different search engine. To avoid erroneous data, we have told our participants to omit questions when they do not feel comfortable to provide us with correct answers. Guideline and questionnaire are available at [2].

D. Questionnaire

In this subsection, we briefly introduce our questions. All questionnaires contained the following questions:

Q1 *Which search engine did you use?*
Q2 *Which search terms did you use?*
Q3 *Which people know the search terms used?*
Q4 *Does the search term itself contain private information?*
Q5 *Which kind of information is on display on the first 20 hits of the search results?*
Q6 *How much information about yourself is displayed?*
Q7 *How old is the information found?*
Q8 *Who has uploaded the information?*
Q9 *Estimate the sensitivity of the information.*
Q10 *Do you approve that this information is available?*

Three new questions have been asked in 2010 and 2011:

Q11 *Have you been surprised to find this information?*
Q12 *Did you allow that this information was published?*
Q13 *What is shown on the images in the search results?*

Depending on the question, we had allowed predefined categories of answers, free-text or a five-point Likert scale. We explain our questions in detail in the next section.

IV. Study

We have obtained 58 questionnaires from 10 participants in 2009, 137 questionnaires from 21 participants in 2010 and 107 questionnaires from 34 participants in 2011. Thus, 65 participants provided us with 302 questionnaires, and each questionnaire contains information about one distinct search result. 152 questionnaires contained one or more questions that have not been answered. 51 participants always answered all questions on each questionnaire. Table I shows the top-5 of questions that have not been answered, and the number of questionnaires where the question has been left unanswered. Note that **Q11** and **Q12** were not part of questionnaires from 2009.

	Question	*Number*
Q6	How much information about yourself is displayed?	123
Q9	Estimate the sensitivity of the information.	65
Q10	Do you approve that this information is available?	60
Q12	Did you allow that this information was published?	49
Q11	Were you surprised to find this information?	44

Table I
TOP-5 OF THE QUESTIONS THAT HAVE REMAINED UNANSWERED

A1: Information Characteristics

How much personal information is available on the Internet?: We have asked our participants to categorize the information that was on display on the first page of the search results or within the first 20 hits (**Q5**). We have provided the following categories: “Memberships” means that the search results indicate that the person concerned is a member of an online community or a social network. The class “Postings” refers to content generated by the person, e.g., a product review on Amazon. “Photos” means that the search result contains pictures showing or made by the searcher. “Locations” refers to places related to the searcher, e.g., the place of living. “Addresses” means telephone numbers, email addresses, etc. “Hobbies” and “Employment” indicates leisure and professional activities, and “Friends” refers to information about social contacts. For the years 2010 and 2011 we have also asked for the content of images in the search results (**Q13**).

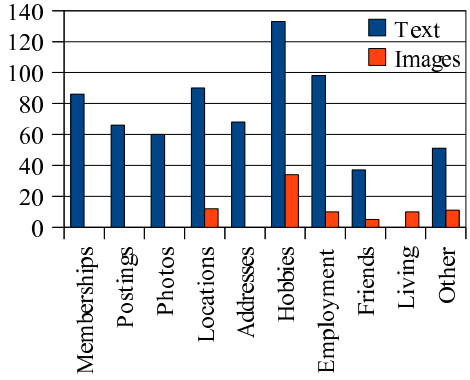

Figure 1. Which and how much personal information is found?

Figure 1 shows that the majority of the textual information available on the Internet refers to hobbies, followed by

Year of the Study	*2009*	*2010*	*2011*
Minimal Age	0.97 (1.67)	1.19 (2.76)	1.54 (2.12)
Average Age	2.07 (1.60)	2.38 (1.85)	3.16 (2.13)
Maximal Age	3.18 (1.80)	3.91 (2.74)	4.65 (2.54)

Table II
AGE OF THE INFORMATION

Uploader	Myself	Friends	Colleagues	Unknown
Number	170	59	101	44
Percent	45%	16%	27%	12%

Table III
UPLOADER

employment, locations and memberships. The distribution of the information categories found on images is similar to the textual results. The information found is well balanced over almost all categories we have provided. Only hobbies seem to be over- and friends underrepresented. Besides our questionnaires, we asked our participants where this information has come from. Important sources of information were student-research papers (recall that our participants were students), web sites of schools and sport clubs that publish awards, placements and team lists, and private homepages.

How old is the information?: Since we were interested to find out if the information found might be out of date, we have asked our participants to write down the range of the age of the information displayed on the first page of the query result (**Q7**). Table II shows the minimal, average and maximal age of the information found, together with the standard deviation (in parentheses). The table shows that, from year to year, the oldest information in the search result gets older. We speculate that publishing personal information regarding our participants at a large scale might have started around 2007, e.g., as a result of online communities like Facebook becoming more and more popular.

Who has made the information available?: From a privacy perspective, it is different if the individual concerned or someone else has uploaded the information. We have asked our participants which category of people might have been responsible for uploading (**Q8**). “Myself” means that our study participant has uploaded the information she has found. “Friends” subsumes friends, acquaintances and relatives. “Colleagues” means that the information has been uploaded with a relation to professional activities, e.g., education, employment or studying. Table III reveals that most of the information our participants have found on the Internet has been uploaded by themselves. Furthermore, a lot of information has been uploaded from colleagues. This observation complements Figure 1, which tells us that “Employment” is the second most-frequent category of information found. A small part of the information has been uploaded by unknown parties.

A2: Search Characteristics

Which search terms have yielded the most information?: Our participants have searched for personal details by using various search terms. In particular, we have encouraged them to search for combinations of the following terms: first name, last name, parts of the postal address of their home and workplace, employment details, email addresses and login names used for instant messaging services or online communities. Note that our participants have filled out questionnaires only for combinations of search terms that have returned at least some personal information. Thus, we have not collected information about inconclusive searches.

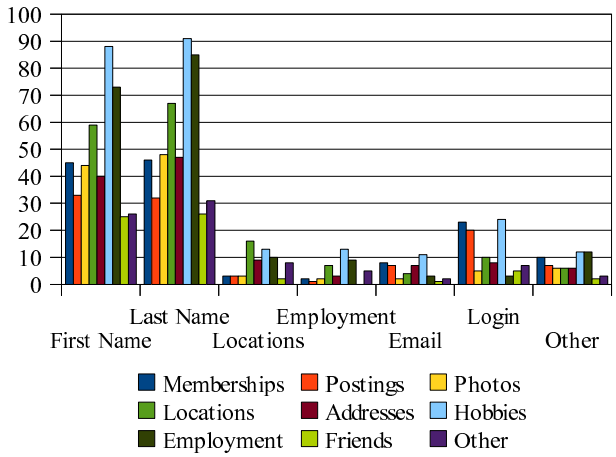

Figure 2. Search terms and information found

Figure 2 shows components of the search term together with the category of information returned. The figure indicates that search terms including at least a part of the real name find most personal details, i.e., user pseudonyms, nicknames, login names etc. play a less important role. However, knowing the login name of a person might be helpful to obtain information related to hobbies, online community memberships and online forum postings which might not be associated with his or her real name.

How much does the search result depend on the search engine?: Our participants were free to use various search engines. To provide a starting point, we have suggested Google.de (general purpose search), Images.Google.de (image search), Yasni.de and 123People.com (person search).

Figure 3 shows which search engines have found which information. In contrast to general-purpose engines, person-search engines should produce more personal results, since they are able to search for semantic information in structured databases, e.g., address registers, indexes of social networks, and in electronic market places like Amazon.com or eBay. Thus, we have expected that person-search engines would be heavily used during our study. However, we have observed the opposite: Provided with the search term “first name last name” (cf. Figure 2), Google has found more information than person-search engines. Only one participant did not find any personal information with Google, but with a person-search engine. The search results of the person-search engines Yasni and 123People were strikingly similar.

 ISBN: 978-1-61208-228-8

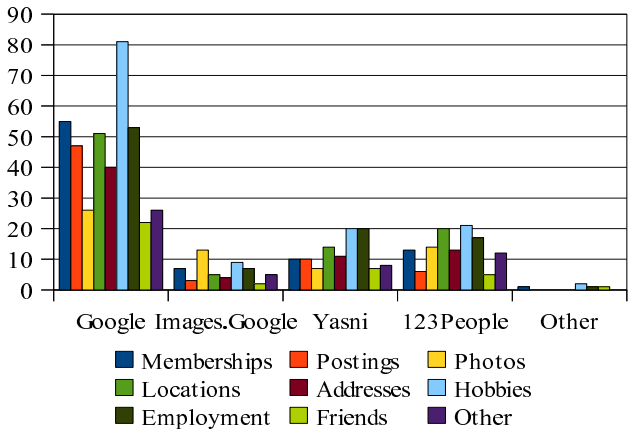

Figure 3. Search engines and information found

A3: Privacy Issues

Have our participants been surprised to find a particular piece of information?: To find out if our participants were able to control which personal information is shown to others, we have asked them if they were surprised by the availability of the information found (**Q11**). This question was answered on 200 questionnaires. Figure 4 shows that, in most cases, the participants found the information they had expected. However, in 20% of all searches our participants were at least surprised by the result, i.e., a significant share of information is available on the Internet without the individuals concerned knowing about it.

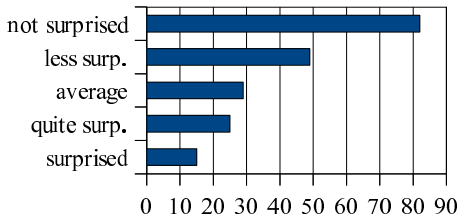

Figure 4. Surprise to find an information

Had our participants given permission to upload the information?: In 2010 we have asked if the information our participants have found has been uploaded with their permission (**Q12**). In 2011, we have refined this question. Now we have asked *which share* of the information has been uploaded with permission. Our participants were asked to regard information they have uploaded themselves as "upload with permission". We have obtained 99 answers to this question in 2010, and 96 in 2011. Table IV shows the results for both years. 15% to 20% of any information found has been uploaded without consent of the individuals concerned. The more detailed results from 2011 indicate that approximately 50% of all searches returned at least a few results where the data has been uploaded without consent. These findings also correspond to Figure 4, where our participants were at least surprised about 20% of the information found.

Year	*With Permission*	*Number*	*Percentage*
2010	no	15	15%
	yes	84	85%
2011	none	20	21%
	few	4	4%
	average	11	11%
	many	13	14%
	all	48	50%

Table IV
UPLOAD WITH PERMISSION

How sensitive do the participants deem the information they have found?: In order to estimate the impact of the information available, we have asked our participants about the sensitivity of the information they have found (**Q9**). This question was answered on 237 questionnaires. As Figure 5 shows, approximately one-fifth of the information found was deemed to be either private or secret, i.e., the participants appraised a significant impact on their privacy.

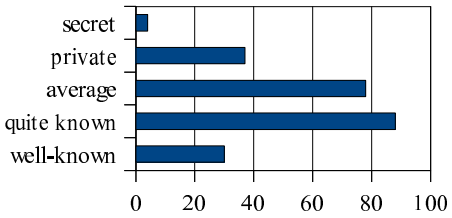

Figure 5. How sensitive is the information?

Do the participants approve that this information is available on the Internet?: As a follow-up question to the last one, we have asked if our participants could tolerate that the information found was on display on the Internet (**Q10**). This question was answered on 228 questionnaires. Since one-fifth of the information has been uploaded without permission, and the same share of information has an impact on the privacy of the participants, we expect that our participants disagree with the availability of at least one-fifth of the information.

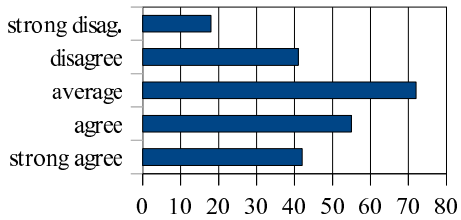

Figure 6. Agreement with availability

As Figure 6 points out, our participants disagree or strongly disagree with the availability of one-fourth of the information. We have calculated the empirical correlation coefficient between the sensitivity of the information (**Q9**)

 ISBN: 978-1-61208-228-8

and the approval of its availability (**Q10**) by regarding the answers to these questions as interval-scaled variables. Both variables are correlated; the correlation coefficient is 0.78. This means that in many (but not in all) cases a participant who thinks that an information is sensitive does not want this information published on the Internet.

Who is able to find which kind of information?: Since the information found depends on the search term, it is important to know who would be able to find which kind of information, i.e., who knows which search term. For example, we know from personal observations that many people do not tell vague acquaintances details about their employment or their place of living, which would enable them to find some information (cf. Figure 2). Figure 7 shows the search terms used together with the categories of people who know these terms. The figure shows that first name and last name are generally known to many categories of people. Locations, employment details, login names and other kinds of identifying information are known to much fewer people. Furthermore, the figure tells us that our participants have shared email addresses and login information with more friends and acquaintances than relatives or other people. We see this as an indication to prevent people like parents or colleagues from learning some information.

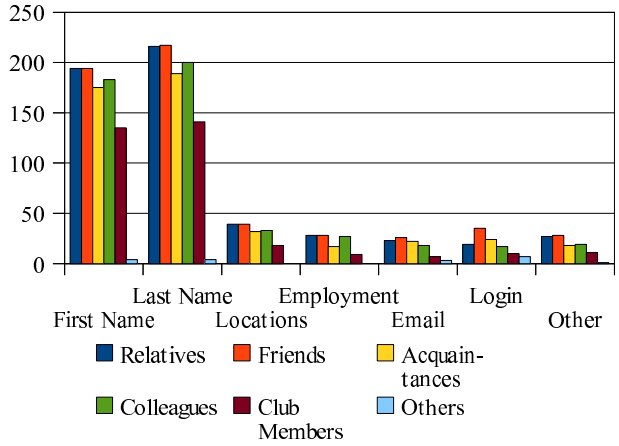

Figure 7. Who knows which search term?

A4: Patterns and Rules

Do groups of individuals with different privacy perception and behavior exist?: In order to design appropriate privacy mechanisms, it is important to identify groups of people with similar attitudes towards personal information on the Internet. Considering that participants have returned different numbers of questionnaires, we have decided for a two-stage procedure: First we apply a clustering approach on all questionnaires. In the next stage, we assign people to each cluster. In particular, we derive a feature vector from each of our 302 questionnaires. The feature vector models the answers of questions about search terms, privacy attitudes and the amount of information available (**Q2**, **Q4**, **Q6**, **Q9**, **Q10**). We have regarded the answers as interval-scaled features where unanswered questions are an additional interval. Since it allows us to inspect clustering results of varying size, we have applied a hierarchical clustering approach. In detail, we have used between-group linkage [16] that starts with one cluster for each feature vector and iteratively combines two clusters with the smallest average distance between all group members in each step. We have used the square Euclidean distance. Due to the hierarchical clustering approach, all questionnaires will be assigned to a cluster. Finally, we have assigned a participant to a cluster if all but at most one questionnaire are a member of the same cluster.

We have manually interpreted cluster sets from 10 to two clusters. According to our interpretation, the most meaningful set consists of four clusters:

Cluster 1: Restrained Publishing This group (105 questionnaires, 18 participants) has found only little information on the Internet, and has not found anything that would have had a severe impact on their privacy: From a privacy perspective, all search results were deemed harmless. The data available has been published with the consent of the individuals. We conclude that this group of people controls very well which information is published on the Internet.

Cluster 2: Incomplete Questionnaires The second group (103 questionnaires, 6 participants) has returned questionnaires that have been filled out incompletely. Because of the anonymity of the study itself, we could not ask the participants for further explanations. We have spent much effort in supervising our participants, and we suppose that they have understood the questionnaire. However, the participants might have found nothing about themselves on the Internet, or they might not have wanted to disclose their results.

Cluster 3: Surprised Individuals from the third group (62 questionnaires, 8 participants) have been negatively surprised about the kind and the extent of personal information they have found about themselves on the Internet. The information has been published without consent of the individuals, or they have published the information without remembering that the data would be available for anybody later on. We assume that this group is less careful in managing their personal data than the first group.

Cluster 4: Generous Publishing This group (32 questionnaires, 3 participants) did find a lot of information about themselves, but does not see this as a problem. The members of this group were not surprised about the kind and extent of the information available. We conclude that this group has a less restrained attitude towards publishing personal information, but manages very well which information may be available to others.

The clustering indicates that many of today's Digital Natives control very well which personal information is published on the Internet. Only eight participants (Cluster 3) were negatively surprised about most search results.

 ISBN: 978-1-61208-228-8

V. Discussion

We were surprised to see that the personal information found on the Internet is well balanced over almost all categories we have provided. Only hobbies seem to be over- and friends underrepresented (cf. Figure 1). Furthermore, we were surprised to see that our participants found nothing unexpected in about 80% of all searches (cf. Figure 4), at least nothing they would deem problematic from a privacy perspective. Nevertheless, we have observed that a certain fraction of information has been uploaded by unknown people and without consent and knowledge of the individuals concerned. Our participants also disagree with the general availability and traceability of some information on the Internet (cf. Figure 6). Furthermore, we have observed that the privacy paradox holds, but to a limited extent: Although most information has been uploaded either by or with consent of the study participant, they disagree with the availability on the Internet of only one fourth of the information.

VI. Conclusion

Due to the advent of social networking sites on the Internet it has become increasingly tempting for individuals to disclose personal details. The potential for misuse of such information is high. To facilitate the design and realization of future privacy approaches, it is important to know the extent and the characteristics of personal data available on the Internet. In this article we have studied which personal information Digital Natives can find about themselves on the Internet. In particular, we have guided 65 undergraduate students of computer science to search for personal information. We have studied the influence of the search engine on the search result, and we have inquired the impact of personal information publicly available on the privacy of the individuals concerned.

As one result, we have gained evidence that Digital Natives are well-adapted to the privacy problems of the Internet. Our explorative study has shown that privacy perception and privacy behavior is different from individual to individual. In particular, we have observed that different search terms return different results, but different search terms are also known to different people. We interpret this as a trend towards managing different digital identities in order to stay in contact with different persons. It might be an interesting topic of research to design privacy approaches that support different digital identities.

Acknowledgement

We like to thank Victoria Kayser for her data analyses.

References

[1] D. M. Boyd and N. B. Ellison, “Social Network Sites: Definition, History, and Scholarship,” *Journal of Computer-Mediated Communication*, vol. 13, no. 1, 2007.

[2] E. Buchmann and K. Böhm, “How Much do Digital Natives Disclose on the Internet – a Privacy Study,” in *Karlsruhe Reports in Informatics 2012/15, ISSN 2190-4782, Karlsruhe Institute of Technology*, 2012. [Online]. Available: http://digbib.ubka.uni-karlsruhe.de/volltexte/1000029029

[3] P. Norberg *et al.*, “The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors,” *Journal of Consumer Affairs*, vol. 41, no. 1, pp. 100–126, 2007.

[4] A. Acquisti, L. John, and G. Loewenstein, “What is Privacy Worth?” in *Proceedings of the 21th Workshop on Information Systems and Economics (WISE’09)*, 2009.

[5] N. F. Awad and M. S. Krishnan, “The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingeness to be Profiled Online for Personalization,” *MIS Quarterly*, vol. 30, 2006.

[6] S. Pötzsch, *Privacy Awareness: A Means to Solve the Privacy Paradox?*, ser. IFIP Advances in Information and Communication Technology, 2009, vol. 298, pp. 226–236.

[7] C. Fuchs, “StudiVZ: Social Networking in the Surveillance Society,” *Ethics and Information Technology*, vol. 12, no. 2, pp. 171–185, 2010.

[8] K. Lewis, J. Kaufman, and N. Christakis, “The Taste for Privacy: An Analysis of College Student Privacy Settings in an Online Social Network,” *Journal of Computer-Mediated Communication*, vol. 14, no. 1, pp. 79–100, 2008.

[9] J.-Y. Son and S. S. Kim, “Internet Users Information Privacy-Protective Responses: a Taxonomy and a Nomological Model,” *MIS Quarterly*, vol. 32, no. 3, pp. 503–529, 2008.

[10] R. K. Chellappa and R. G.Sin, “Personalization versus Privacy: An Empirical Examination of the Online Consumer’s Dilemma,” *Information Technology and Management*, vol. 6, pp. 181–202, 2005.

[11] Universität St. Gallen, “Abschlussbericht des Projekts Sicherheit vs. Privatheit - Vertrauensfaktoren im Umgang mit Daten und Konsequenzen für die digitale Identität,” http://isprat.net/isprat-projekte, 2011.

[12] M. Kurt, “Determination of in Internet Privacy Behaviours of Students,” *Procedia - Social and Behavioral Sciences*, vol. 9, no. 0, pp. 1244 – 1250, 2010.

[13] I. Oomen and R. Leenes, *Privacy Risk Perceptions and Privacy Protection Strategies*, ser. IFIP International Federation for Information Processing, 2008, vol. 261, pp. 121–138.

[14] K. B. Sheehan, “Toward a Typology of Internet Users and Online Privacy Concerns,” *The Information Society*, vol. 18:21/32, 2002.

[15] L. Sweeney, “k-Anonymity: A Model for Protecting Privacy,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[16] L. Rokach, “A Survey of Clustering Algorithms,” in *The Data Mining and Knowledge Discovery Handbook*, 2nd ed., L. Rokach and O. Maimon, Eds., 2010, ch. 14, pp. 269–298.

 ISBN: 978-1-61208-228-8

Issues and Risks Associated with Cryptocurrencies such as Bitcoin

Félix Brezo and Pablo G. Bringas
Avenida de las Universidades 24, 48007
DeustoTech Computing (S3lab), Universidad de Deusto
Bilbao (Bizkaia), España
Email: {felix.brezo, pablo.garcia.bringas}@deusto.es

***Abstract*—Bitcoin is an electronic currency designed to use a public protocol that implements it in a totally decentralized manner, so as not to need the control of any central issuing organization that manages it. Though still in development, it has been proven to be a modern payment system referred to have been used in some procedures commonly associated to money laundering or trafficking of illegal substances of various kinds. Thus, in this article, we analyse those features which transform such a cryptocurrency in a useful tool to perform any kind of transactions far from the control of any kind of regulatory agency, as well as we pinpoint some of the fields in which their usage can derive in new illicit behaviours.**

Keywords-bitcoin; cryptoanarchism; cryptocurrency; fraud; speculation; virtual money.

I. Introduction

The phenomenon of virtual coins is neither new nor particularly recent. The inception of the concept is usually located in some distribution lists in the mid nineties [1]. Thus, nowadays, there are lots of virtual currencies being used with more or less success in the Internet: from Pecunix to e-gold, passing from all kind of virtual currencies associated to leisure applications such as Second Life's Linden Dollars, whose economy has also been studied by some authors [2], [3], and the well-known Facebook's Facebook Credits.

The novelty of Bitcoin —originally created by Satoshi Makamoto in 2009 [4]— is the fact of being a public protocol implementing a *peer-to-peer*-based cryptocurrency. Its definition as a distributed currency comes from the absence of an existing central entity in charge of regulating either the value or the amount of the total number of existing coins. It is the network itself, making use of the computational capacity of its own users, the one that manages and maintains it. This calculation capacity is used, amongst other things, to manage the transaction history —what is known as the *block chain*— and to confirm and validate each and every new transactions to be happening in the future.

The characteristics of its distribution protocol are providing Bitcoin with a strong boost in certain communities in the internet. In fact, the value of the total transactions performed annually has raised in 2011 to 150 million dollars as stated by the website blockchain.info while the protocol manages peaks of 25,000, 30,000 and even 47,000 daily transactions as displayed on Figure 1.

Thereby, the remainder of this paper is structured as follows. Section II states some of the key aspects that make Bitcoin special. Section III defines some scenarios in which the use of this cryptocurrency may lead to illegal behaviours. Section IV summarises and defines the main conclusions to be extracted about this new reality.

II. Bitcoin key features

Against this background, the appearance of a currency with such special characteristics leads to a new scenario which possibilities have never been explored before. In this section, we collect some of the aspects that make this cryptocurrency a differentiating factor.

A. Distribution protocol

Its specific characteristics stablish a particular method for allocating bitcoins amongst those nodes in the peer-to-peer network that share computing capacity for the distributed maintenance of the system. Notwithstanding, the coin assignation process effectiveness is inversely proportional to the total network's capacity. By protocol, the maximum number of total monetary units is set to the amount of 21 million bitcoins [4] —sometimes represented using the Thai baht (ISO 4217 code: THB) symbol ฿ or, in text format, as BTC—, being reached approximately by 2040. This definition of the protocol makes possible to predict the total number of coins in circulation in any moment of the history. For instance, by mid-2012, this figure rounds the 9 million of already distributed coins, with an estimated value in the markets of 52.99 million dollars.

By June 2012, the distribution process takes place approximately every 10 minutes. With that frequence, ฿50 are assigned randomly to one of the nodes that have contributed to solve a given mathematical problem: finding the result of a given *hash*. A hash function is a computable function that takes as an input a set of elements (usually strings) and maps them in a finite output range, typically fixed-length strings, being theoretically impossible to reverse it. Thus, the only way of finding the input of a given hash is to bruteforce the calculations. The complexity of this problem is adjusted every two weeks to maintain the rhythm of 6 handouts per hour. To illustrate this situation, in the current circumstances, the great amount of connected

 ISBN: 978-1-61208-228-8

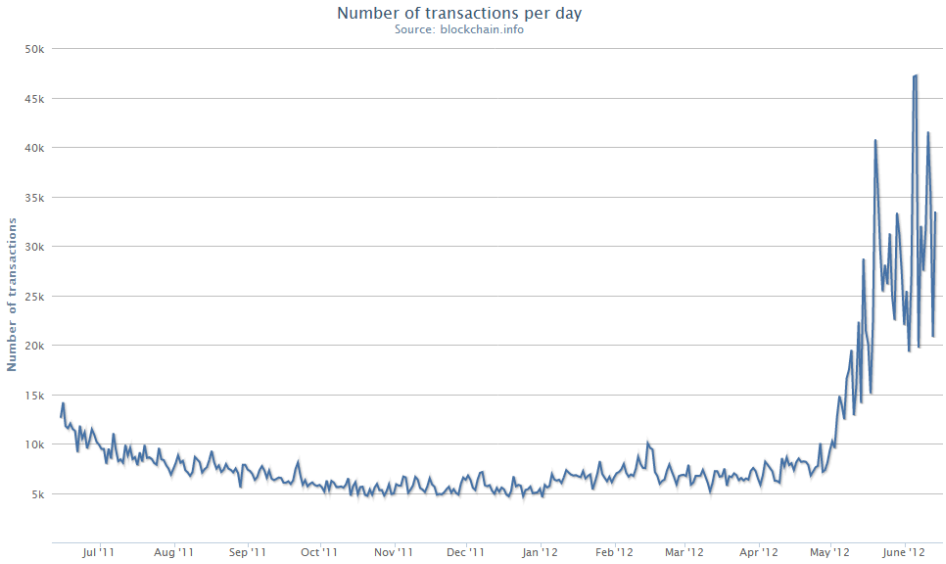

Figure 1. Total daily Bitcoin transactions as stored by Blockchain website.

Table I
ESTIMATED TIME NEEDED TO OBTAIN ฿50 DEPENDING ON THE COMPUTING CAPACITY DEPLOYED BY A GIVEN USER, MEASURED IN MILLIONS OF HASHES PER SECOND.

Computing capacity	Average time needed	Daily generation
5 *MHashes*/s	43 years, 45 days	฿0.0032
20 *MHashes*/s	10 years, 285 days	฿0.0127
100 *MHashes*/s	2 years, 57 days	฿0.0635
500 *MHashes*/s	157 days, 9 hours	฿0.3177
2,000 *MHashes*/s	39 days, 8 hours	฿1.2706
10,000 *MHashes*/s	7 days, 20 hours	฿6.3532
50,000 *MHashes*/s	1 days, 13 hours	฿31.7661

devices and the increasing capabilities of the network to work out this hashes will let a standard user exploiting not optimized hardware —i3 processor, 4 cores, NVIDIA GT240 graphic card— and contributing without any kind of affiliation to the development of the network —practice known as *solo mining*— to develop a computing capacity of 20 million hashes solved per second. In other words, this standard user would receive the mentioned ฿50 per block approximately every 11 years. Although we could also use the processors computing capacity, this chance is deprecated as it is estimated to be between 30 and 50 times slower that the one produced by modern graphic cards designed for GPU computing —8 Intel ® Core™ i7 processors are hardly capable of reaching the figure of 2 million hashes/second—. In the table I, we show some estimations of the needed time for the obtention of the ฿50 depending on the computing capacity deployed by the user.

This reality has led to the concept of *pool mining* versus the already described of *solo mining*. The philosophy of the former lies in the grouping of isolated users computer power under one unique operator with the final goal of hoarding a much greater power to receive the bitcoins assignations more continuously and relying less on luck. Thus, pool operators retain a fixed amount —between the 2 and the 10 per cent depending on the mining pool and its distribution method— in concept of management and maintenance of the website, resharing the remaining amount proportionally to the computer power provided by all the users who have trusted on that operator. Usually, the corresponding traffic runs on port 8,332, but, as a side note, it is true that there exist some pools that encapsulate Bitcoin traffic through port 80 to avoid the blocking of certain firewalls.

B. Currency exchange markets

There are numerous ways of buying or selling bitcoins. In fact, in Bitcoincharts [5], more than 50 active markets were already listed in June 2012, with very different available currencies as shown in Figure 2, but the options are unlimited: from accessing to exchange websites in which perform Paypal, Liberty Reserve, WebMoney or OKPay transferences or executing directly bank transfers depending on the site (some of the most well-known are MTGox, Bitcoinmarket or BTC-e); till the acquisition of the gift vouchers to be used in sites such as eBay, Amazon or Steam; passing by buying directly any kind of goods and services in E-Commerce platforms supporting Bitcoin such as osCommerce, which counts with a plugin that permits the transferences using Bitcoin since May 2011 [6].

C. Off-line payments and transactions acceptance period

One of the main problems of using a cryptocurrency is the need to verify that the coin has not already been previously used by its owner to perform another transaction. Bitcoin

 ISBN: 978-1-61208-228-8

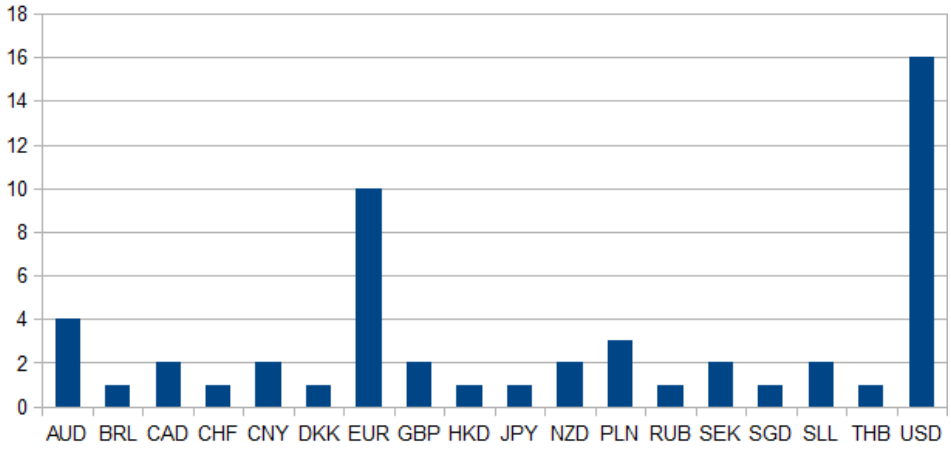

Figure 2. Number of available markets depending on the exchange currency.

choose a distributed checking model instead of checking the legitimacy of its use against a single point.

Thus, this verification process requires the disclosure of the transfer to the rest of the network nodes, which are responsible for verifying that the given coin has not been used in more than one occasion by introducing duplication. To achieve this objective, it is necessary that the charging platforms were connected, being impossible to confirm whether a transfer is valid or not if this were not so. This process often lasts an undefined period of time ranging from 10 to 60 minutes [4]. However, this period is potentially reducible if it includes the payment of a small commission to be assigned to that node in the network that is able to verify the transfer. This fact introduces a new motivation component for user sharing their computing power whenever the total amount of bitcoins have been handed out.

D. Transactions tracking

Anyway, the anonymity of the transactions is not guaranteed by the protocol itself since this is not its ultimate goal [4]. In fact, the need for open verification requires the disclosure and publication of every transfer. Recent studies have used traffic analysis tools so as to identify the called *egocentric networks*, showing that anonymity is not a standard pattern in the majority of the transactions [7]:

- In the analysis conducted by Reid et al. [7], a collection of transactions were monitored from an *self-incriminated* thief, to whom the authors assume the interest of remaining anonymous so as not to be tracked. The point is that the authors were capable of identify relationships of up to $n = 1, 2, 3$ hops between the victim and the alleged cybercriminal.
- Another case brought by the authors referred to the announcement of Wikileaks to accept *anonymous donations via Bitcoin* using a single public key per transaction. The authors call into question the real anonymity of the donor.

In short, it is assumed that the strength of Bitcoin does not rely on boostening the anonimity: in fact, the transactions will be only as anonymous as impossible it were to uniquely identify the connection of the user that runs them.

III. Prospective on criminal behaviour

The aforementioned characteristics can be exploited in a healthy and reliable way, but also agglutinate features that may be exploited in the shadow. In this section, we will cover some of them.

A. Speculative movements sensitivity

As most economic elements, bitcoins are also sensitive to radical changes linked to speculation. The exchange of bitcoins for euros, dollars, pounds or any other cryptocurrency is particularly sensitive to speculative movements and the subjective perception of its value. Figure 3 shows the evolution of the exchange markets since the birth of the cryptocurrency till the actual situation. As can be seen, the perception of their value has markedly fluctuated to reach the astonishing levels of 25 dollars/bitcoin, figures far removed from the far more stabilized 5.3 dollars/bitcoin in June 2012; identifying around June 2011, what some authors have defined as the *Bitcoin Bubble* [8].

Speculation is also present in the long run. Considering a constantly growing user population over time and knowing that the total amount of bitcoins ends up being a finite number, the value of ฿1 will tend to represent goods with increasing value: that is to say, Bitcoin will move towards becoming an economy with hints of finally evolving into a potentially deflationary system. The main operational problem that entails the divisibility of money, is already tackled by Bitcoin as each coin is divisible up to 10^{-8} parts. Or what would be the same, each coin could be divided in parts equivalent to μ฿10 (10 *nanobitcoins*), placing the total amount of independent coins in 2,100 trillion units.

 ISBN: 978-1-61208-228-8

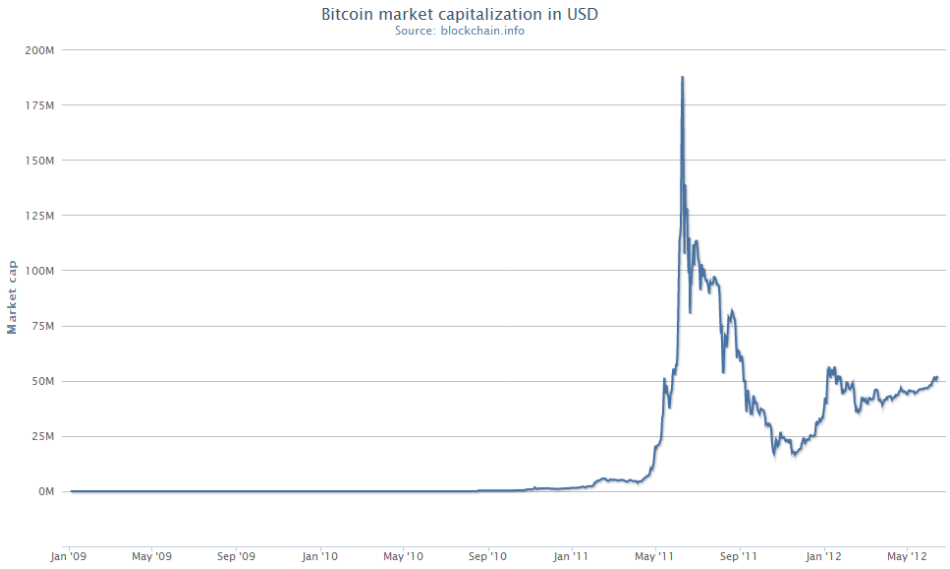

Figure 3. Evolution of the Bitcoin market capitalization in USD since the inception of the cryptocurrency.

However, the actual circulation of all coins is an utopia. The reason for this is the existence of hundreds of coins which are assigned daily to users who lost or will lose the control over them for reasons that do not need to be associated with fraudulent practices:

- It could happen that bitcoins get either lost or forgotten somewhere inside a legitimate computer.
- This hardware, can also be affected by any kind of technical issues that may make impossible the recovery of the cyberwallets.
- It could also happen that the legitimate users would forget the password that unlocks them.

B. Money laundering and illegal traffic

Using online entertainment platforms that somewhat simulate an internal monetary system to hide certain cyberdelictive behaviours is not a recent issue. Mallada [9] described how certain criminals have started to exploit the possibilities offered by some *Role-Playing Games* (RPG) such as *World of Warcraft* or the virtual world of *Second Life*. For instance, in the case of the latter, the author estimates its Gross Domestic Product in $500 million.

Taking into account that Bitcoin develops a payment system which final objective is no longer to simulate but to substitute traditional currencies, the implications and chances of the aforementioned potential applications are even bigger. Even more, despite the characteristics in the previous subsection which do not make the payment method intrinsically anonymous, it is true that the performance of transfers could end being untraceable if some appropriate preventive measures were taken [10]:

- The performance of transferences behind a *proxy* or other anonymisation systems such as networks like Tor. This is the case of the *Silk Road* online merchant [11]: to access to the offered services it is compulsory to be using the Tor browser so as to preserver the anonymity of both, buyers and sellers. Already referenced by some authors as the *Amazon of illegal drugs* [12], [13], in some countries such as the USA some policy initiatives have been carried out to close similar platforms to prevent the illegal drug trafficking [14].
- The execution of n transferences at a time towards an account that, later and as part of a *batch* or *offline* self-controlled, redistributed the received money in m different accounts. Given a great number of transactions from a great number of users, it would be virtually impossible to track the coins from their real owner lost in a woods of relations. This process, as defined in Nakamoto's original description of the system [4], can take place under the following circumstances: either by means of individuals offering the chance of voluntarily increase the anonymity, or by means of organized networks offering this and other similar services in a sort of *dark-market* operating with cyberwallets in the cloud, also kwnown as eWallets. These eWallets are Bitcoin accounts created and managed by third parties which offer to their users an easier way of keeping in touch with their bitcoins. However, given the irreversibility of the transactions, there is no real guarantee that those transferences performed to or from an eWallet will be run, apart from the confidence that the user has on the service itself. A possible scenario may be the following:
  1) A big amount of bitcoins is to be sent to a new eWallet anonymously created using, for example, tools like Tor, described above.
  2) Small transfers are made from such wallet in the

 ISBN: 978-1-61208-228-8

cloud to a set of similar eWallets —inside or outside the original service— before receiving in a collection of new Bitcoin accounts the money to operate with them again.

This approach raises points that are likely to require an automated account management, suggesting a greater knowledge and better use of the computer tools already available.

C. Virtual pickpocketing

System features make the *coins* being stored as .dat files in the computers of their owners. Those may make backup copies to be stored in alternative computers or devices —or even in the *cloud*— in order to maintain access to them whatever kind of failure takes place as this would imply the automatic loss of the control on the coins as there would be no way to recover them.

Nonetheless, any user or program with logical access to the .dat file could execute transactions on his/her/its own if this file has not been adequately protected. Although the use of the so called *military* encryption standards —as, for example, AES-256, the Advanced Encryption Standard specification for encrypting electronic data adopted by the U.S. government [15], widely considered as a *de facto* cryptographic standard— is accessible by any user since the protocols are public and implemented in numerous libraries and open-source platforms; it is a standard pattern coming across with unprotected files, resulting in a stream of malware threats uniquely dedicated to stealing these files.

In this line, Laboratories such as Kaspersky Labs, deactivated in late March the Command & Control channel (simplifying, the C&C channel is the tool for managing and controlling a network of kidnapped computers or *botnets*, used in a bunch of different criminal ways) of Hlux/Kelihos. In this case, *Hlux* was a botnet specialised in stealing Bitcoin wallets [16], proving that even cybercriminals have found a new battlefield from which obtain illegal benefits easy to monetize. Meanwhile, the Sections on ÇyberIntelligence and Criminal Intelligence of the Federal Bureau of Investigation (FBI) stated in a recently filtered intelligence report [17] that there exist Zeus samples —a Zbot trojan version specialised in the recruitment of machines as part of a botnet— specially designed to operate in the shadow using the victim's resources in the dark.

D. Unauthorized use of computing power

However, the use of similar networks may also have other purposes. Similarly to what happens with the distributed computing projects, could be used downtime CPU or GPU of infected machines for the generation process of exchange without the user's knowledge. Currently, mining of bitcoins is not considered a profitable business if it is not performed using the appropriate hardware. In many cases, the revenues produced in the mining process are not able to cover the energy costs of the generation.

An important aspect to avoid the detection of any unauthorized process is to give the user the fewer indicator as possible. This also occurs when hiding illegal mining processes. In this sense, the use of *applets* which make use of processor idle time —for example, assigning the mining processes the lowest possible priority as the *Bitcoinplus* miner does— or not exploiting all the capabilities of the graphics card —avoiding overloading and warming— would become essential. This may be enough to avoid the end inexpert user of the machine noticing the symptoms of a machine being remotely used, unless he/she manually checks the use percentage of the processor.

In the scenario of a middle-size-botnet compounded by 10,000 usable computers, and considering a computing power of an average of the 0.005 bitcoins/day and machine, the derivated exploitation would round the 50 bitcoins/day—about 200 euros/day in the most common exchange markets—. Taking into consideration the dismantling in 2010 of the botnet Butterfly, on a combined action of Panda Labs along with the Spanish Guardia Civil, and which size stood at 13,000,000 infected machines, the numbers speak for themselves about the potential benefit in which may incur the administrator of such a network. Being able to exploit these characteristics would make any botmaster leverage from a seemingly innocuous part of the computing power of the infected machines obtaining a direct benefit and increasing artificially the electric bill of the legitimate user.

E. A likely jump to the real world

For the writing of this report, the authors have independently contacted with certain European business that publicly state that accept Bitcoin as a payment and which actually use this cryptocurrency as bait to attract Bitcoin new customers. In the public list available online in Bitcoin's official Wiki [18], where the reader can mainly find local hotels and restaurants among other businesses, it becomes clear that the use of this currency, which in principle might seem limited to Internet shopping, has surpassed, although still timidly, the barriers to real world.

Its presence introduces an additional component of complexity in the control of buying and selling process as there is no official entity to recognize its use and, therefore, there is a full absence of any kind of official change in the different central banks. What is more, this may lead to problems associated to tax evasion and fraud because of the impossibility of calculating the official value of the transaction in what can be considered, technically, a barter economy: that is to say, exchanging goods or services for bitcoins, which may be defined under certain circumstances as *other cryptographic goods with some subjective value*.

 ISBN: 978-1-61208-228-8

F. Alternative cryptocurrencies based on Bitcoin

Bitcoin's source code can be widely studied since it is distributed under the MIT License. This, almost naturally, has led to various *sister* cryptocurrencies which implement specific characteristics that differentiate them from the original:

- IXcoin (IXC). This is a cryptocurrency with a parallel development to Bitcoin but scheduled to have a shorter maturity period, as the maximum number of coins (also 21 million) will have been generated by 2015. This suggests that speculative movements in the maturation period of the currency are potentially more violent and less predictable.
- Devcoin. This is a currency that allocates the 90% of the resources generated to developers participating in open source programs so as to fund their work, reserving only the remaining 10% to the miners.
- Namecoin (NMC). Also based on the architecture of Bitcoin, the namecoins constitute a currency which target is creating a domain name system (DNS) using the .bit TLD. The objective is to provide resources and tools to protect the community against a censorship to be potentially applied by a central entity (like ICANN in the case of domain names). Thus, the .bit domains are maintained entirely by a Bitcoin-like peer-to-peer network.

In another development, the philosophy Bitcoin uses to prevent double spending and limit the proliferation of fraudulent currency has been studied by some authors as Becker et al. [19] for a generalization and export to other areas that share similar needs, even beyond the domains management Namecoin tries to implement.

IV. Conclusion

Though yet in an underground development phase and, mostly, pretty unknown for the general public, the proliferation of these new payment alternatives brings many uncertainties. What is more, the intrinsic complexity of the protocol and the necessity of having some relatively advanced knowledge on cryptography and computer studies to understand its real behaviour, make these cryptocurrencies the perfect place for speculation and misinformation. For instance, as already stated, there is a widespread belief of the mere fact of using it is sufficient guarantee to perform anonymous transactions, when this is not true by definition.

At the same time, the absence of a regulatory central organism and the chance of not being able to fix the prizes in a explicit way as it happens with the traditional cryptocurrencies, defines a new scenario on an economy strictly ruled by the market movements with all the consequences that this fact leads to in terms of control of massive speculative efforts. At this point, amongst the possible failure scenarios the most urgent for Bitcoin, excepting a dramatical reduction of users which may devaluate the currency once mature, is, precisely, a global governmental campaign against its use.

Just before the end, we can conclude that there is a real risk of a recurrent illegal use of the cryptocurrency. The great number of existing markets and the possibility of exchanging easily bitcoins by euros, pounds or dollars, make this new method the perfect vehicle to perform every kind of transactions related to money laundering or illegal traffic of substances, with all the legal implications associated to the jurisdictional limitation of the criminal acts performed in the cyberspace.

Acknowledgments

The authors would like to thank the Spanish Guardia Civil from their interest and support to this research and, personally, to the person of José María Blanco.

References

[1] D. Barok, Bitcoin: censorship–resistant currency and domain system for the people, in *Networked Media*, Piet Zwart Institute, 2011.

[2] P. Ernstberger, Linden dollar and virtual monetary policy, in *Macroeconomics*, p. 20, Department of Economics, Economics I, Bayreuth University, 2009.

[3] B. Mennecke, W. D. Terando, D. Janvrin, and W. Dilla, It's just a game, or is it? Real money, real income, and real taxes in virtual worlds, in *Communications of the Association for Information Systems*, volume 20, pp. 134–141, 2007.

[4] S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, http://www.bitcoin.org, 2009.

[5] B. Charts, *All Markets*, 2012, available online http://bitcoincharts.com/markets/list/ [retrieved: August, 2012)].

[6] *Bitcoin payment module*, 2011, available online http://addons.oscommerce.com/info/8007/ [retrieved: August, 2012)].

[7] F. Reid and M. Harrigan, An analysis of anonymity in the Bitcoin system, in *Privacy, Security, Risk and Trust (PASSAT), 2011 IEEE Third International Conference on and 2011 IEEE Third International Confernece on Social Computing (SocialCom)*, pp. 1318–1326, IEEE, 2011.

[8] B. M. Maurer, Money nutters, in *Economic Sociology - The European Electronic Newsletter*, p. 5, 2011.

[9] C. Mallada, Las nuevas tecnologías y el blanqueo de Capitales: Second Life, entretenimiento online y el método delictivo, in *Challenges and Opportunities of Online Entertainment. Proceedings of the 8th International Conference on Internet, Law and Politics*, pp. 199–209, Universitat Obverta de Catalunya and HUYGENS Editorial, 2012.

[10] S. Martins and Y. Yang, Introduction to bitcoins: a pseudo-anonymous electronic currency system, in *Proceedings of the 2011 Conference of the Center for Advanced Studies on Collaborative Research*, pp. 349–350, IBM Corp., 2011.

 ISBN: 978-1-61208-228-8

[11] N. Christin, Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace, CMU-CyLab-12-018, 2012, Id: paper.tex 1286 2012-07-30 21:29:14Z nicolasc.

[12] E. Jacobs, Bitcoin: A Bit Too Far?, in *Internet Banking and Commerce Vol.*, volume 12, 2011.

[13] M. J. Barrat, SilkRoad: Ebay for drugs, in *Addiction*, volume 107, pp. 683–683, Wiley Online Library, 2012.

[14] J. Manchin, *Manchin Urges Federal Law Enforcement to Shut Down Online Black Market for Illegal Drugs*, 2011, Press Release http://manchin.senate.gov/public/index.cfm/2011/6/manchin-urges-federal-law-enforcement-to-shut-down-online-black-market-for-illegal-drugs [retrieved: August, 2012].

[15] Biryukov, A. and Khovratovich, D. and Nikolić, I., *Advances in Cryptology-CRYPTO 2009* , 231 (2009).

[16] Kaspersky Lab, How Kaspersky Lab and CrowdStrike Dismantled the Second Hlux/Kelihos Botnet: Success Story, 2012.

[17] Directorate of Intelligence: Cyber Intelligence Section and Criminal Intelligence Section, Bitcoin Virtual Currency: Unique Features Present Distinct Challenges for Deterring Illicit Activity, in *FBI Intellligence Assesment*.

[18] *Real world shops*, 2012, available online https://en.bitcoin.it/wiki/Real_world_shops [retrieved: August, 2012)].

[19] J. Becker, D. Breuker, T. Heide, J. Holler, H. Rauer, and R. Böhme, Can We Afford Integrity by Proof-of-Work? Scenarios Inspired by the Bitcoin Currency, in *Workshop on the Economics of Information Security WEIS 2012*, 2012, JEL Classification: E42, Q30.

 ISBN: 978-1-61208-228-8

Knowledge Sharing through Social Networks

Liaquat Hossain
Complex Systems and Project Management Program
The University of Sydney, NSW Australia
Department of Informatics, Lund University, Sweden
Liaquat.hossain@sydney.edu.au or @ics.lu.se

Simon Reay Atkinson
Complex Systems and Project Management Program
The University of Sydney, NSW Australia
Simon.reayatkinson@sydney.edu.au

Rolf T Wigand
Departments of Information Science and Management
University of Arkansas, Little Rock
rtwigand@ualr.edu

Sven Carlsson
Department of Informatics
Lund University, Sweden
Sven.carlsson@ics.lu.se

Abstract—**We suggest that a lack of understanding of the social and human factors in the design process may lead to the failure of knowledge sharing or KS in most organizations. The basis for KS in organizations is embedded in participants' action and experience. We propose that successful KS initiatives require (1) attention to communication patterns of individuals or groups working in different divisions of an organization and (2) the development of IT systems that support both strong and weak ties between participants. Specifically, a distinction between different network structures as they relate to the concept of structural holes is drawn for highlighting how types of network structures effect sharing of explicit and tacit knowledge. Additionally, strong and weak tie theories are applied to develop a framework for potential IT-based initiatives aimed at addressing structural holes of communication. A set of propositions is proposed with their implication for designing KS systems in organizations. This paper concludes that sociological perspective in achieving a balance between the different types of ties (i.e., strong and weak ties) could assist in the maintenance and ongoing creation of new knowledge without having the networks to be redundant.**

Keywords- knowledge sharing; transfer systems; social networks; tie diversity; structural hole.

I. INTRODUCTION

Social networks or SN is an interdisciplinary behavioral approach for the study of human actors, their relationships and interdependencies. SN draws theoretical and methodological foundations from areas such as communication, sociology, social psychology, psychiatry, organizational science and graph theory. It involves theorizing, model building, and empirical studies, which focuses on uncovering patterns of communication among actors, organizations, states etc. Therefore, uncovering the communication practices among members from different organization, which would result in detecting cliques, isolates and brokers and their role in the network forms the basis for application of SN to study the knowledge sharing in organizations.

Social ties, a critical aspect for understanding the formation of cliques, isolates and brokers and their role in the network can be seen from two perspectives: (i) strong ties, and (ii) weak ties. Social network literature suggests strong ties or close and frequent interactions among people from different organizations are likely to lead to redundant information because they tend to occur among a small group of actors in which everyone knows what the others know. Contrary to strong ties, weak ties or distance and infrequent relationships provides access to novel information by bridging the disconnected groups and individuals in organizations [7, 46]. The usefulness of strong or weak relationships among people in different organization for effective coordination needs further investigation and especially in the context of knowledge sharing [3, 4, 5]. Furthermore, the relationships among strong and weak ties and computer mediation for successful knowledge sharing context have yet to be explored [7, 8, 9]. Also, the use of SNA and its outcome in improving communication flow and its process as a measure for developing effective knowledge sharing systems is yet to be explored.

The transfer of knowledge between individuals, groups, communities or systems is regarded as Knowledge Sharing (KS) in organizations [1, 2, 3]. Research suggests that understanding social interaction between different interest groups within an organization is a critical component of effective KS [4, 5, 6]. The relevance of a social interaction approach has been debated in areas such as systems design, organizational process redesign, process improvement and artificial intelligence *per se* [7, 8, 9]. The social approach for designing KS is based on the argument that communication between individuals, teams, groups, and communities is critical to the development and sustainability of a knowledge-creating organization [10,

 ISBN: 978-1-61208-228-8

11, 12, 13]. It is also suggested in studies that an appropriate structure of KS is essential for facilitating effective sharing in organizations [1]. Specifically, we suggest that KS is dependent on the structure of the social communication network at play in an organization.

Organizational science literature also highlights that communication network structure provides insight about the communication patterns of individuals working in an organization [7]. Therefore, an understanding of the communication network at play needs to be viewed as an essential part of the design of KS systems in organizations. There is also a tension in the organizational structure, strategy and process literature when applied to KS in organizations. Studies suggest that organizations should not start with structure but with a task-and-person based foundation that incorporates both authority and responsibility [14]. Therefore, the design of the KS structure should be based on the study of the existing communication structure. Communication networks may suggest how individuals, groups, communities or systems interact in an organization and can be used as a basis for KS process of an organization [1, 2, 3].

In this paper, we first provide a background to organization as a network o people. A person or a group of people united for some purpose is considered to be a form of organization [15, 16]. Cyert and March [15] suggest that organization needs to be viewed as a form of coalition. That is, an organization is considered to be a coalition of individuals, some of them organized into sub-coalitions. Arrow [16] highlights that formal organizations, firms, labor unions, universities, or government, are not the only types of entities that represent the term 'organization'. For example, the market system has elaborated methods for communication and collective decision-making and, therefore, can be interpreted as an organization [16].

It is further suggested by Mintzberg [17] that organization structures have both a formal and informal structure. Formal organizational structure is usually represented by the organization chart and widely accessible by the internal and external members. It is also suggested in the organizational science literature that every organization is a network of people [15, 18, 19, 20]. An analysis of the communication network can help us in understanding the information exchange, patterns, coalition and power of the individual members in an organization [7, 21]. The distinction between formal and informal organization structure can be drawn by looking at the types of interactions, or links, between individuals or agents in an organization. For example, the legitimate network refers to formal structure and the shadow network refers to the informal structure of an organization [20]. In the legitimate network, interactions or links are either (i) formally and intentionally established by the powerful members of the organization or (ii) established well-understood, implicit guiding principles, which is accepted by the members of the organization [20]. On the other hand, the shadow network consists of links that are spontaneously and informally established by the individuals among themselves during the interaction process in the legitimate system [20].

It is also evident that the shadow system does not coincide with the rigid boundaries of the legitimate system. Shadow system is classified to have porous boundaries and therefore considered to the principal route for interaction between individual agents in an organization or in an inter-organizational network [20]. We argue that the KS system needs to be designed by conducting a thorough requirement analysis of both the legitimate and shadow network. This is important as the legitimate network may provide a normative view of how individuals should share knowledge while the communication network analysis of shadow network will assist KS system designers in understanding the descriptive view of individual agents' communication patterns. This information could later be used to directly address issues of *structural holes* that may or may not exist in an organization or in a department [20].

In this paper, we first highlight that successful KS initiatives require (1) attention to communication patterns of individuals or groups working in different divisions of an organization and (2) the development of IT systems that support both strong and weak ties between participants. In particular, we provide a distinction between different network structures as they relate to the concept of structural holes. We also highlight how types of network structures effect sharing of explicit and tacit knowledge. Additionally, strong and weak tie theories are applied to develop a framework for potential IT-based initiatives aimed at addressing structural holes of communication. A set of propositions is proposed with their implication for designing KS systems in organizations.

II. DESIGNING KNOWLEDGE SHARING SYSTEMS

KS systems design evolved from the traditional structured systems design literature. Scientists, engineers, technicians, and programmers initially performed the design of technology-based systems in the 1950s and 1960s [22, 23]. Kling highlights that design flaws were the major impeding factor for ensuring the optimal use of computer-based information systems in organizations [23, 24, 25, 26]. The design of computer systems for supporting collaborative work requires careful attention in five key areas--planning, analysis, design, implementation and support [27, 28]. The importance of careful examination of these phases for ensuring the success of systems implementation has been addressed in organizational design literature as well [17, 29, 30].

 ISBN: 978-1-61208-228-8

Previous studies suggest that systems design is essentially a social process [23, 24, 25, 26]. Therefore, the social role of systems analyst is one of the critical success factors for the successful design and implementation of the system [1, 31, 32, 33]. This social role is essential for the collection of relevant information from different disciplines and people during the requirements analysis phase of the systems development [23]. Therefore, the design of technology-based products has to be in line with social and organizational dynamics [23]. In fact, there is a danger of systems failure or not receiving high rate of user acceptance if systems design issues are considered separately from the organizational issues. This is a common problem for the implementation of multi-module software systems such as enterprise resource planning [23].

Studies suggest that the social systems design approach by Mauro Mauro Design Inc. improved the performance of the New York Stock Exchange trading systems. The systems analyst from Mauro Mauro Design Inc observed the traders at the Stock Exchange for six months prior to start coding new software together with 30 iterations in testing their new systems [34]. Kling and Star [34] highlights that analyses that cover the complexity of social organization and the technical state of the art is critical to the design or use of human-centered computing. This analysis can provide the systems designers with insights both the technological characteristics of a computerized system and the social arrangements under which the system will be used [35].

For example, it is highlighted in studies that understanding the distinction between the legitimate and shadow network structure is an important first step towards the design of the KS in organizations. Understanding the shadow network structure requires a communication network analysis so that the patterns of exchange between agents in a network can be understood. It is highlighted in previous studies that a successful knowledge creation process requires an established communication network. Communications network structure deals with individual communication pattern in an organization or in a unit of work. KS design can be viewed as a social process because it requires interaction between all parties moving through developmental phases together in order to produce a system that is efficient and effective. It creates ownership in a system, which alleviates many of the problems traditionally associated with implementing a new system, resistance to change, resistance to imposed authority, training, etc. This serves as a basis for the development of a conceptual model of KS in organizations [36].

Nonaka [37] and Brown and Duguid [38] also support that knowledge creation is essentially a social process. It is suggested in case studies such as Nucor Steel and Buckman Labrotaries [39] that understanding the interaction of individuals, groups, teams and communities in knowledge networks leads to a successful KS in organizations. Therefore, it can be seen that there is a growing interest among social scientists to view KS as socially constructed and embedded in social networks and communities of practice [11]. These findings clearly highlight that the dichotomy of KS systems design can be seen from three perspectives—technological determinism, systems rationalism, and socio design. In the following section, we provide a brief overview of the systems design literature as it relates to the design of KS in organization.

III. SOCIAL DESIGN OF KNOWLEDGE SHARING SYSTEMS

Social design refers to joint design of both the technological characteristics of a system and the social arrangements under which it will be used [35, 40, 41]. Bijker [40] argued that the development of technological systems should be viewed as a social process, not an autonomous occurrence where relevant social groups will be the carriers of that process. Kling et al [23] further highlights that these social choices are considered to be an integral part of computerization, even though they are not formally decided or completely within the control of any one person. For example, company A is adopting portable computers so that they can improve the flexibility of people's work situations and relationships. However, company A still insists that their employees report to work daily during the regular working hours. Therefore, employees of company A have very little flexibility to work from a remote location even though they have access to the technology infrastructure provided by the company. In contrast, the underlying operational philosophy of company B is to allow its employees to work from remote location so that it provides maximum flexibility and optimal use of portable computing. This example illustrate that it is not only the technology that guide successful operation, but the guiding principles or social design of work practices that organizations decides to pursue.

Managers, therefore, must address the cultural side of change when implementing software systems such as ERP as it increases fear among managers that the availability of company wide information may challenge their authority [42]. The biggest impediment to knowledge transfer is corporate culture and the biggest difficulty in managing KS is changing people's behaviour. Therefore, organizations need IT infrastructure to make progress or to provide the facilitation of knowledge networks, but the use of ICT for managing KS activities should be supported by introducing proper organizational processes, people and content. There is also a growing interest in considering a social network approach to understand the KS design in organizations. Social network analysis refers to the method of analysing social structures and relational aspects of structures that exist in a communication network. It is highlighted in the previous section that communication network structure

 ISBN: 978-1-61208-228-8

can be viewed as a legitimate or shadow network. That is, an organization's structure may suggest how the legitimate communications network should work and the shadow network structure may suggest how the communication flow occurs at an organization. Therefore, social network analysis is continuing to play a significant role in developing a deeper understanding of the actual process of communication flow between individuals [18]. Additionally, one can conclude from all of the above that social network analysis has the potential to play a significant role in the design and implementation of knowledge management systems.

IV. BUILDING INFO-CULTURE FOR KS

Socialization, externalization, internalization, and combination can be seen as a mechanism for the creation of knowledge in organizations [37]. Here, the socialization in organization refers to the conversion of tacit knowledge to new tacit knowledge through social interaction and shared experiences [3]. The combination mode deals with the creation of new explicit knowledge by merging, categorizing, reclassifying and synthesizing existing explicit knowledge. Both the externalization and internalization mode refers to interactions and conversation between tacit and explicit knowledge where externalization deals with the conversion of tacit knowledge to new explicit knowledge and internalization deals with the creation of new tacit knowledge from explicit knowledge. Table 1 provides an overview of four modes of knowledge creation.

TABLE 1. FOUR MODES OF KNOWLEDGE CREATION

Modes of Knowledge Creation	**Characteristics**	**Examples**
Socialization	Conversion of tacit to new tacit knowledge	Apprenticeship, user training
Combination	Creation of new explicit knowledge	Survey reports
Externalization	Conversion of tacit to new explicit knowledge	Lessons learned
Internalization	Creation of new tacit from explicit knowledge	Learning and understanding from reading and discussion

It can be seen from Table 1 that knowledge sharing and creation is dependent on the modes of knowledge creation. Here, socialization is seen as an important aspect for the conversion of tacit knowledge into new tacit knowledge. For example, the development of an "*infoculture*" is the first essential step for creating knowledge-based organizations [39]. The study of Nucor Steel highlights that three essential elements—superior human capital, high-powered incentives and a high degree of empowerment guide the knowledge creation process [39]. Nucor Steel used a *group-based incentive* mechanism to encourage people to start sharing knowledge that in fact lead to the development of an *infoculture* in their organization. This incentive mechanism was introduced at all levels of the organization so that Nucor could only reward group-based performance.

Social network analysis is increasingly used to develop a better understanding of the shadow system network that is considered to be a true representation of the communication patterns that exist in an organization. Social networks can be defined as an individual's relations and contacts with others [21, 42]. Social network analysis can be seen as a method that allows us to analyse social structures and relational aspects of the structures that exists in a communication network between individuals, teams, groups and communities. The argument advanced in this paper is that once the IT-based KS systems put into practice or implemented in organization, it becomes a social network. Therefore, the social design of this IT-based KS should be established through a thorough analysis of both the formal and informal social networks that may exist in an organization. It is argued here that the design of IT-based KS should be able to accommodate the facilitation of the communication patterns or flow process that exists in a department or in an organization. Therefore, the social dimensions of KS can be described from two perspectives—the first is the role of socialization and community building as a backbone social infrastructure for KS, and the second is the IT-based KS systems. IT-based KS systems are also considered as social systems as this KS systems link people as well as machines.

Wellman [43] suggests that computer supported social networks help sustain strong, intermediate and weak ties which provide information and social support in both a specialized and broad-based relationships. It is also important to note that there are direct and indirect ties exist between agents or the participating agents engaged in KS. It is clear that these ties are embedded in both the legitimate and shadow network of an organization. This, when combined with what is known about computer supported social networks mentioned above, may provide valuable insights for the effective design of IT-based knowledge management systems. This is discussed further in the following two sections.

A. KS through Strenghts of Ties

It is indicated earlier that organizations can be viewed as a network of people. In particular, we discussed two types of networks—legitimate and shadow and its implications for the design and sustainability of KS provided. These networks consist of individuals working

in an organization and can be seen as redundant or non-redundant. A structural hole is referred to as a relationship of non-redundancy between two or more contacts [21]. Non-redundant contact between individuals can be seen as disconnected either directly or indirectly. Here, the disconnected direct non-redundant contacts suggest that there is no direct contact with one another and the indirect contacts suggest that one has contacts that exclude the others. Burt [21] further suggests that the two contacts provide network benefits as a result of the structural holes. Here, we discuss the concept of structural holes together with the strong and weak ties metaphor as it relates to KS systems design in organizations.

Burt [21] suggests that two criteria—cohesion and structural equivalence can be used as an indicator for detecting structural holes. Cohesion criterion refers to direct connection between the contacts. For example, two contacts A and B are redundant to the extent that a strong tie connects both A and B. Here, this strong tie between contacts A and B indicates the absence of structural holes (e.g. the relationship between father and son, or people who frequently connects with each other for social occasions). However, structural equivalence concerns indirect connection by mutual contact. For example, both A and B are structurally equivalent to the extent if they both have same contacts.

This nature of the contacts between the executives and persons in their network is referred to as strength of ties [44]. Intense, emotion-laden, and reciprocal relationships that require time and energy to create and maintain can be a reflection of strong ties. Weak ties on the other hand, reflect loose networks and are best explained by the concept of a bridge [44]. The strength of the tie has traditionally been viewed as bearing on the overall amount and content of information associated with the contact. It is however suggested in previous studies that novel and non-redundant information is available through weak ties more than through strong ties [44, 45]. Strong ties can be seen as advantageous because they allow for quick flow of information and social support. Furthermore, strong ties are reliable, easily available, and important when dealing with conflicts, crises, and uncertainty [46].

Granovetter's [44] theory of strong and weak ties highlights the importance of weak ties in providing information. A weak tie is defined as a "casual acquaintance" and a strong tie is a formal relationship defined by a high-shared knowledge base and multiple interactions [45, 46, 47]. Burt [21] further suggests that weak ties provide a useful mechanism for understanding the strength of structural holes in a communications network. We believe that both these types of ties offer unique opportunities for developing a theoretical base for the design of KS systems in organizations from both a theoretical and an applied perspective. Studies suggest that weak relationships such as casual acquaintances, do not take as much time and effort to cultivate as friendships or community of practice. It is therefore easy to have more acquaintances than friends. A larger number of acquaintances can provide access to information about more out-groups. Most importantly, acquaintances offer the potential for (1) a relationship that takes limited time and effort and (2) offers the most potential for non-redundant and, thus, valuable information and knowledge. Specifically, Burt [21] proposed a direct relationship between the number of structural holes and the rate of return on player's investment in terms of time and energy and social capital (Figure 1). Here, the shape of the curve is to indicate the general relationship between structural holes and human capital as opposed to any validated and specific function.

B. Supporting KS through Social Capital

We propose above that the value of social interaction and social exchanges needs to be taken into consideration by a designer of an IT-based KS system. We further propose here that the design of an effective IT-based KS system should allow for an economic use of time and energy in the growth of social capital.

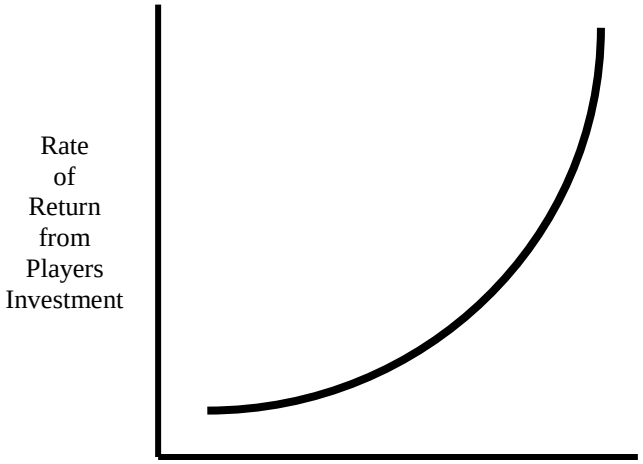

Figure 1. Relationship between Rate of Return and Structural Holes [21]

To discuss the specific effects of an IT-based KS system on social capital we will illustrate how one can increase capital by acting on the following five t propositions:

1. Maximizing weak ties in one's network increases the potential for innovation and/or market penetration [44].
2. IT-based KS systems are an effective means for establishing and maintaining weak ties [43].

 ISBN: 978-1-61208-228-8

3. Maximizing the number of structural holes in one's network increases the potential for innovation and/or penetration [21].
4. A finite number of strong ties can be maintained [21].
5. Minimizing the number of strong ties allows for more allocation of resources to the application and creation of new knowledge.

Assume "Entrepreneur A" (E.A.) has lived and worked in city "Home" for a number of years. During this time, E.A. has established a number of strong ties due numerous in-person exchanges (Figure 2). E.A. soon realizes that both new ideas and/or potential markets for his product have become too redundant and that a larger network is required for further growth. E.A. decides to explore some potential new contacts via various forms of IT (Figure 2).

After establishing a number of loose contacts, E.A. decides to strategically strengthen ties with those who offer the greatest number of resources. The inherent communication barriers associated with numerous forms of IT motivate E.A. to invest time and energy into more face-to-face interactions with one of the selected new contacts. E.A. spends weeks (or, perhaps, longer) on-site with new contact E.B. who resides in Home2 and begins to strengthen their tie. E.A. also spends times with some of E.B.'s contacts and establishes a number of new relationships. During this time, E.A. maintains his ties back home by using IT channels. The maintenance of these ties requires (1) relatively little effort given the established shared knowledge base and (2) a strategic approach as to which contacts from Home make the most "economic" sense to maintain. E.A. maintains other new contacts by using the same methods used while at home. E.A. takes special care to maintain a weak tie, or no tie, between E.B. and other new contacts (Figure 3).

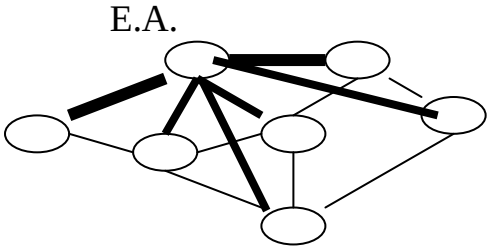

Figure 2. Social Network with proximal, strong ties

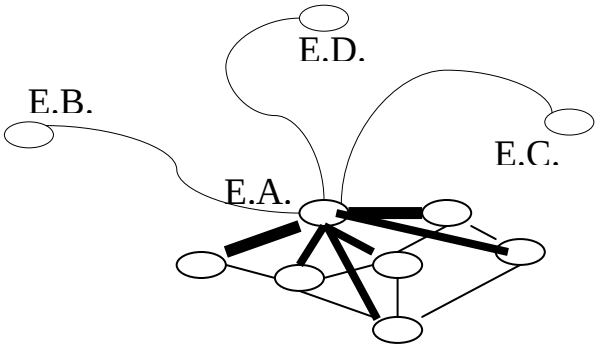

Figure 3: Social Network with proximal, strong ties, and new "distal", weak ties

Once the required shared knowledge base is established, E.A decides to move to Home3 to establish a stronger tie with E.C. E.A. maintains strong ties with Home and E.B. through efficient use of IT. Most importantly, E.A. maintains the structural holes between Home, E.B., E.C., and E.D. (Figure 4).

This process continues until E.A. reaches a maximum number of strong ties that can be maintained without having to "fill" in, or bridge, important structural holes due to a limited number of strong ties that can be maintained. E.A. eventually shows a significant amount of growth in social capital (see Figure 3) and invests more time in managing the flow of knowledge and information rather than actively searching for more capital (Figures 5 and 6).

One can better understand the demands of a system designer by combining an understanding of cognitive demands and limitations with the social behaviors and needs of end users like E.A. That is, for example, E.A. potentially realizes his cognitive capacity with the development and maintenance of 6 strong ties and the knowledge and information flow resulting from the increase in social capital. With every developed strong tie the demand imposed by the maintenance and utilization of a growing—and, thus, more complex—social network increases. Therefore, the ultimate requirement of an IT-based KS system is to minimize cognitive demand in the maximization and utilization of social capital. From this perspective, the IT-based KS system complements the individual and increases the potential for innovation and knowledge creation.

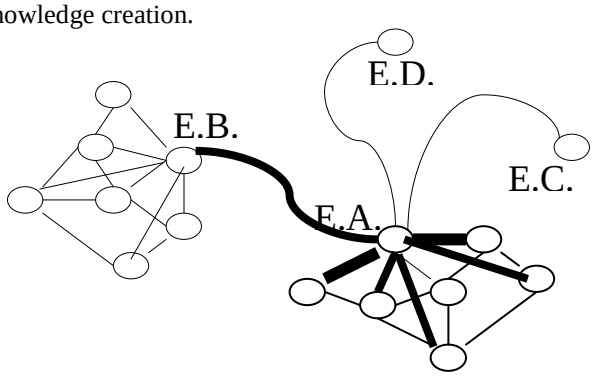

Figure 4. Example of Social Network with proximal, strong ties, distal weak ties, and a new distal strong tie

 ISBN: 978-1-61208-228-8

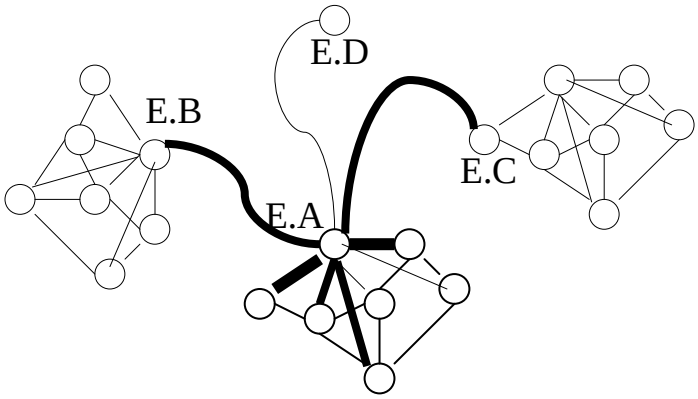

Figure 5. Example of Social Network with proximal strong ties, new distal weak ties and strong ties

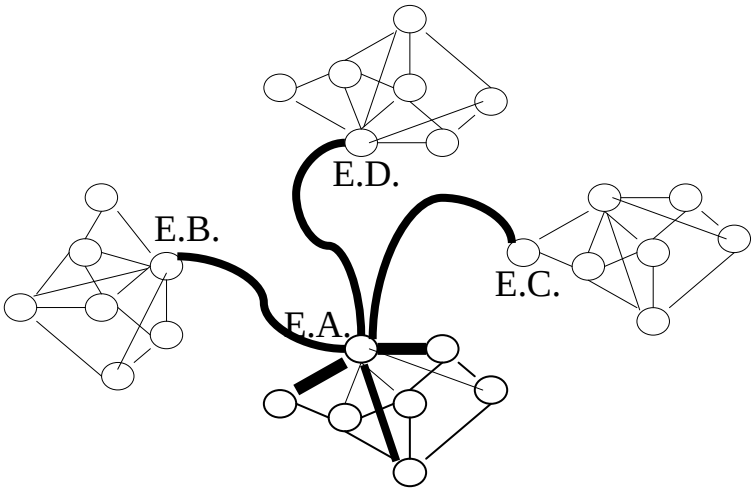

Figure 6. Example of Social Network with the same number of strong ties as shown in Figure 4, but with a significant increase in the amount of social capital

V. CONCLUSIONS

It is concluded that complex systems design such as knowledge sharing systems requires careful attention to the social context for which is systems is developed. Therefore, we suggest here that the designer of KS systems is required to develop the systems based on observations of the social interaction and social exchange needs of the individuals working in an organization. We further highlight that the design of KS should be based on an understanding of both the legitimate and shadow network structure of an organization. We recommend that social network analysis is a useful methodological paradigm for the purpose of eliciting the communication patterns for understanding the shadow network structure. It is suggested here that structural holes provide network benefits for developing social capital among participating members in knowledge sharing and conclude that strong and weak tie is a useful metaphor for fitting the structural holes. In conclusion, the following propositions can be used to develop a better understanding about how to effectively and efficiently leverage information technology for the development of social capital which help support knowledge sharing in organisations

1. Maximizing weak ties in one's network increases the potential for innovation and/or market penetration [44].
2. IT-based KS systems are an effective means for establishing and maintaining weak ties [43].
3. Maximizing the number of structural holes in one's network increases the potential for innovation and/or penetration [21].
4. A finite number of strong ties can be maintained [21].
5. Minimizing the number of strong ties allows for more allocation of resources to the application and creation of new knowledge.

VI. REFERENCES

[1] Davenport, T. H., and Prusak, L. (1998). *Working Knowledge: How organizations manage what they know*, Boston, Massachusetts: Harvard Business School Press.

[2] Hansen, M. T., and Von, O. B. (2001). 'Introducing T-shaped managers: Knowledge management's net generation,' *Harvard Business Review*, March, pp. 38-49.

[3] Alavi, M., and Leidner, D. (2001). 'Review: Knowledge *Management and Knowledge Management Systems: Conceptual Foundations and Research Issues,' MIS Quarterly, 25, 1, 107-36.*

[4] Pickering, J. M., and King, J. L. (1995). 'Hardwiring Weak Ties: Inter organizational Computer-Mediated Communication, Occupational Communities, and Organizational Change,' *Organization Science*, 6, 4, 479-486.

 ISBN: 978-1-61208-228-8

[5] Ivari, J., and Linger, H. (1999). 'Knowledge work as collaborative work: A situated activity theory view,' in *Proceedings of the Thirty-Second Hawaii International Conference on Systems Sciences*, IEEE Computer Society Press, Los Alamitos, CA.

[6] Tuomi, I. (1999). 'Data is more than knowledge: Implications of the reversed hierarchy for knowledge management and organizational memory,' in *Proceedings of the Thirty-Second Hawaii International Conference on Systems Sciences*, IEEE Computer Society Press, Los Alamitos, CA.

[7] Wigand, R. T. (1988). 'Communication Network Analysis: History and Overview,' In: *Handbook of Organizational Communication*, G. Goldhaber and G. Barnett, Ablex Publishing Corporation, Norwood, New Jersey, 319-359.

[8] Nonoka, I., and Takeuchi, H. (1995). *The Knowledge-Creating Company: How Japanese companies create the dynamics of Innovation*, New York: Oxford University Press, Oxford.

[9] Earl, M. (2001). 'Knowledge management strategies: Toward a taxonomy,' *Journal of Management Information Systems*, 18, 1, 215-233.

[10] Argyris, C., and Schon, D. (1978). *Organizational Learning: A Theory of Action Approach*, Reading, MA: Addison Wesley.

[11] Kogut, B., and Zander, U. (1992). 'Knowledge of the firm, combinative capabilities, and the replication of technology, *Organization Science*, 3, 3, 383-397.

[12] Watson, R. T. (1999). *Data Management: databases and Organizations* (2nd ed.), John Wiley, New York.

[13] Levine, L. (2001). 'Integrating knowledge and processes in a learning organization,' *Information Systems Journal*, Vol. 18, No. 1, pp.21-33.

[14] Drucker, P. (1974). *Management: Tasks, Responsibilities, Practices*, Butterworth-Heinemann Ltd. Oxford.

[15] Cyert, R. M., and March, J. G. (1963). *A Behavioral Theory of the Firm*, Prentice-Hall Inc., NJ, USA.

[16] Arrow, K. (1974). *The Limits of Organization*, Norton and Company, New York.

[17] Mintzberg, H. (1979). The structuring of organizations: a synthesis of the research, Prentice-Hall Englewood Cliffs, N.J.

[18] Charan, R. (1991). 'How networks reshape organization-for results,' *Harvard Business Review*, Sept.-Oct., 12, 479-494.

[19] Nohria, N., and Eccles, R. G. (1992). *Networks and organizations*, Cambridge, MA: Harvard University Press.

[20] Stacey, R. D. (1996). *Complexity and Creativity in Organizations*, Berrett-Koehler Publishers, San Francisco.

[21] Burt, R.S. (1992). *Structural Holes: The Social Structure of Competition*, Cambridge, Massachusetts: Harvard University Press.

[22] Yourdon, E. (1989). *Modern Structured Analysis*, Englewood Cliffs, N.J: Yourdon Press.

[23] Kling, R., Crawford, H., Rosenbaum, H., Sawyer, S., and Weisband, S. (1998). *Learning from social informatics: Information and communication technologies in human context*, Center for Social Informatics, Indiana University, IN.

[24] Kling, R. (1991). 'Computerization and social transformations,' *Science, Technology and Human Values*, Vol. 16, No. 3, pp. 342-367.

[25] Kling, R. (1993). 'Organizational analysis in computer science,' *The Information Society*, Vol. 9, No. 2, pp. 71-87.

[26] Kling, R., and Jewett, T. (1994). 'The social design of worklife with computers and networks: an open natural systems perspective,' *Advances in Computers*, Vol. 39, Marshall Yovits, Ed., Orlando, FI: Academic Press, pp. 239-293.

[27] Avison, D., and Shah, H. (1997). *The Information Systems Development Life Cycle*, The McGraw-Hill Co. Ltd.

[28] Whitten, J. L., and Bentley, L. D. (1998). *Systems Analysis and Design and Methods*, Irwin McGraw-Hill. Fourth Edition.

[29] Schon, D. A. (1967). *Technology and change: the new Heraclitus*, Delacorte Press, New York.

[30] Mintzberg, H. (1989). *Mintzberg on Management: Inside our strange world of organizations*, Free Press: New York.

[31] Forsythe, D. (1992). 'Blaming the user in medical informatics,' *Knowledge and Society: The anthropology of science and technology*, Vol. 9, pp. 95-111.

[32] Forsythe, D. (1994). 'Engineering knowledge: the construction of knowledge in artificial intelligence,' *Social studies of science*, Vol. 24, pp. 105-113.

[33] Suchman, L. (1996). 'Supporting articulation work,' in *Computerization and Controversy: Value conflicts and social choices*, 2nd Ed., Kling, R., (Ed.), San Diego, CA: Academic Press.

[34] Kling, R., and Star, S. L. (1998). 'Human-centered systems in the perspective of organizational and social informatics,' *Computers and Society*, March, Vol. 28, No. 1, pp. 22-29.

[35] Kling, R. (1987). 'Computerization as an ongoing social and political process,' in *Computers and Democracy: A Scandinavian Challenge*, Bjerknes, G., Ehn, P., and Kyng, M. (Ed.), Vermont: Gower Pub. Co.

[36] Armbrecth, F. M., and Chapas, R. B. (2001). 'Knowledge management in research and development,' *Research Technology Management*, V. 44, No. 4, pp. 28-48.

[37] Nonaka, I. (1994). 'A dynamic theory of organizational knowledge creation,' *Organization Science*, 5, 10, 14-37. [38] Brown, J. S., and Duguid, P. (1998). 'Organizing knowledge,' *California Management Review*, 40, 3, pp. 90-111.

[38] Brown, J. S., and Duguid, P. (1998). 'Organizing knowledge,' *California Management Review*, 40, 3, pp. 90-111.

[39] Gupta, A., and Govindrajan, V. (2000). 'Knowledge flows within multinational corporations,' *Strategic Management Journal*, 21, pp. 473-496.

[40] Bijker, W. E. (1997). *Of Bicycles, Bakelities, and Bulbs: Toward a theory of sociotechnical change*, MIT Press.

[41] Bijker, W. E., and Law, J. (1992). *Shaping technology building society: studies in sociotechnical change* (Eds.), MIT Press.

[42] BarNir, A., and Smith, K. A. (2202). 'Interfirm alliances in the small business: The role of social networks,' *Journal of Small Business Management*, Milwaukee, Vol. 40, Issue 3, pp. 219-232.

[43] Wellman, B. (1996). 'For a social network analysis of computer networks: A sociological perspective on collaborative work and virtual community', *Proceedings of the 1996 conference on ACM SIGPR/SIGMIS conference*, Denver, Co. pp. 1-11.

[44] Granovetter, M. (1973). 'The strength of weak ties,' *American Journal of Sociology*, Vol. 78, pp.1360-1380.

[45] Granovetter, M. S. (1974). *Getting a job: A Study of Contacts and Careers*, The University of Chicago Press, Ltd., London.

[46] Granovetter, M. S. (1982). 'The strength of weak ties: A network theory revisited,' In: *Structure and network analysis*, edited by Peter V. Marsden and Nan Lin, Sage Publications, pp. 1054.

[47] Hansen, M. T. (1999). 'The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge Across Organization Subunits,' *Administrative Science Quarterly*, 44, 1, 82-111.

 ISBN: 978-1-61208-228-8

Evaluation of e-Government Policies: Overlooked Aspect of Public Interest

Dalibor Stanimirovic, Mirko Vintar
Institute for Informatization of Administration
University of Ljubljana, Faculty of Administration
Ljubljana, Slovenia
dalibor.stanimirovic@fu.uni-lj.si, mirko.vintar@fu.uni-lj.si

***Abstract*—The majority of existing methodologies for evaluation of e-government policies is underdeveloped and partial, preventing comprehensive and objective evaluation. This situation consequently results in poor quality of planning and implementation process, while further diminishing positive effects and decreasing public consumption of e-government services. One of the most frequently overlooked aspects of e-government policies evaluation is the concept of public interest, which is not given sufficient attention within existing evaluation methodologies, reducing the legitimacy of policy making in the field. The significance of public interest is often elusively defined, while its dimensions are somehow rendered particularly within the financial benefits. Paper provides an analysis of more than 50 methodologies for evaluation of e-government policies, exploring the presence of public interest aspect within. Analysis offers an insight into the current evaluation practice enabling detection of its deficiencies as well as their mitigation, and could facilitate a significant contribution to more evidence-based evaluation of e-government policies.**

Keywords-e-government policy; evaluation methodology; public interest

I. INTRODUCTION

Various studies indicate that further e-government development is one of the most important factors of public sector rationalization [1][2][3], as well as faster countries' development [4][5][6]. Despite considerable investments in e-government in recent years (European Union (EU) countries are investing approximately 2.2% of Gross Domestic Product (GDP) in Information and Communication Technology (ICT) in the field of public sector [7][8][9]) the expected effects in terms of reducing costs and increasing the effectiveness of public sector are still rather ambiguous, while user acceptance of e-government services is far below government anticipations. Disclosed issues and present public finance situation along with increasingly stringent austerity measures require careful direction of further e-government investments, particularly focusing on proficient evaluation of e-government policies and their effects – be it on national, local or sectoral level.

Past experience in the field necessitates the development of methodologies for evaluation of e-government policies (the collective term "methodologies" is used denoting approaches, studies, indicator models, measurement frameworks and similar undertakings for evaluation of e-government policies) which could enable e-government decision-makers to conduct more qualified and quantified preparation, execution and evaluation of e-government policies including their broader societal implications. Considering e-government development so far, we have been witnessing a big gap between supply and demand of public e-services in most countries, which can be prevailingly attributed to "politically driven" development rather than "evidence based" evaluation and selection of e-government policies [10][11][12].

Notwithstanding the increasing number of different evaluation methodologies emerging in the last years [13][14][15], some aspects of e-government policies evaluation have been largely disregarded, particularly public interest as one of the foundations of public policy making. This rather unsatisfactory state of affairs has led to growing calls for a reassessment and rebalancing of the rationalisations in evaluation of e-government policies, and in particular for a greater weight to be given to the public interest [16].

The paper discusses the main features of existing methodologies for evaluation of e-government policies and analyses their evaluation foci trying to establish the extent to which existing methodologies facilitate evaluation of public interest. Deriving from the aforementioned research objectives the paper is focusing primarily on the following interrelated research questions:

1) Overview and study of existing methodologies for evaluation of e-government policies and summary of their characteristics.
2) Analysis of the existing methodologies for evaluation of e-government policies regarding the aspect of public interest.

From the methodological point of view, research represents a typical in-depth analysis, while research activities are embedded in two-phase incremental methodological framework. Combining different techniques of qualitative research methods [17], the initial part of the study has focused on the analysis of primary and secondary

 ISBN: 978-1-61208-228-8

sources, whereas deriving from obtained research results, the conclusive part of the research is striving to integrate theoretical and practical aspects regarding the research subject. Selection of research methods was adapted to the research field [17][18] given the complexity of e-government evaluation initiatives.

Following the introduction, the second section of the paper presents the concept of public interest and an overview of the relevant literature while outlining various directions in evaluation of e-government policies and related issues and barriers. The third section provides an analysis of existing methodologies for evaluation of e-government policies and summary of their characteristics. The fourth section explores presence of the public interest aspect in the existing methodologies and sketches the research findings. The last section contains the review of the overall research, discussion on its limitations and subsequently submits the final arguments and observations regarding the research results and future work.

II. CONCEPT OF PUBLIC INTEREST AND METHODOLOGIES FOR EVALUATION OF E-GOVERNMENT POLICIES – STATE OF THE ART

The public interest is not a unitary concept: different public interests are relevant in different scenarios and need to be weighted differently depending on the circumstances [16]. Public policy makers, expert public and citizens have to find the appropriate balance of the numerous public interests that may exist in any given situation. Finding this balance will not only involve comparing the relative importance of one public interest to another but also involve the contemplation of the interconnection of public interests and considering the broader impact that these may have in turn on other public interest networks [19].

Although there are different conceptions of public interest inaugurating different research perspectives, most frequently, public interest is viewed as a set of substantive ideals against which all policy proposals should be judged. The concept is thus recognized as a symbol to which all agree (few people are opposed to the public interest) and to which special interests appeal in order to rationalize their policy desires [20]. Regarding the contextual platform consisted of political and cultural framework of a particular society and the economic resources at its disposal, the public interest is defined as the aggregate of the fundamental goals that the society seeks to achieve for all of its members. Thus, the art of government consists of achieving a harmonious rather than a destructive balance among conflicting goals [21]. Some authors have studied the public interest concept from the aspect of substantive truths or principles. These truths or principals are not formal tests that any public policy must meet; however general they may be and however much skill may be required to apply them in particular cases, they provide substantive guidance to the proper content of public policy [22]. While others consider public interest and its dimensions primarily as a process of public action, primarily bargaining and competition between different interest groups, resulting in the overall social consensus [23]. Summarizing numerous and occasionally complex definitions of the public interest concept, it can be generally regarded as a set of commonly agreed goals arising from the inclusive and transparent decision-making procedure, based on compromise and shared values of well-defined social community.

Despite its complexity and diversity of research approaches and perspectives, public sector concept is elaborated relatively well in theory, whereas it has failed to gain significant attention in majority of e-government evaluation undertakings. Consideration of public interest concept in methodologies for evaluation of e-government policies is often inadequate and superficial covering prevailingly financial factors and omitting all other societal aspects while reducing the potential of evaluation process for objective and comprehensive evaluation.

A. *Methodologies for evaluation of e-government policies*

In parallel to e-government development there have emerged numerous methodologies, trying to evaluate its development and effects on different parameters of government operation. According to their characteristics and subject of evaluation these methodologies could be classified in typical groups presented below.

1) *Front-office maturity and readiness*

The most known benchmark measurements on the EU level have been conducted by Capgemini [7][8][9], while the most renowned benchmark measurements on the global scale have been carried out by the UN [14][24], Accenture [25] and Brown University [26]. While focusing primarily on web site analysis, indicators from these methodologies are not precise enough to ensure comprehensive evaluation and validation of e-government policies on the national level (see critical analysis of such benchmark measurements from [12][27][28], etc.). Some other important benchmark measurements dealing with e-readiness are: The Global Information Technology Report [15], Digital economy rankings [29] and United Nations e-Government Survey [14][24]. These benchmark measurements deploy different sets of indicators for benchmarking e-readiness and information society in general. Being predominantly focused on front-office change and infrastructural requirements, these rather extensive methodologies are hardly providing evaluation of public interest, incapacitating its incorporation in the process of further e-government development.

2) *Effects and impacts of e-government policies*

Within a number of methodologies focusing on ex-ante and ex-post evaluation of e-government policies we could highlight: MAREVA [30], eGEP [31], WiBe 4.0 [32] and AGIMO [33]. MAREVA is dealing with ex-ante and ex-post evaluations of e-government policies on the basis of parameters such as profitability, risks, benefits, necessity; similar aspects are evaluated by WiBe 4.0. AGIMO and eGEP are striving to identify and analyse costs of establishment, provision, maintenance and performance as well as impacts of e-government services. In general we could find these methodologies very exhaustive in terms of the large number of indicators; however they rarely address the concept of public interest comprehensively, while some

 ISBN: 978-1-61208-228-8

particularised dimensions of public interest are normally amalgamated with financial benefits.

Considering significant organizational changes in the public sector organizations, induced by implementation of e-government policies, the research in the field is primarily focused on joined-up e-government model [34], organizational changes in the direction of network government [35], management and external factors which affect e-government development [36], business process change, information management capacity and organizational capabilities [37][38] and organizational culture [39]. An overview of related methodologies reveals there is no clear consensus on organizational changes caused by e-government implementation, and consequently no comprehensible methodology to measure implications of transformed public sector organizations for the various social aspects and implementation of public interest.

3) National-level development

Surveys often highlight political and sociological factors as the most important external factors affecting e-government development. This aspect is partially discussed in United Nations e-Government Survey [14][24] through indicators such as e-participation, e-inclusion, e-consultations, and other research in the field [9][25][40] is focusing on political and sociological indicators such as accessibility, digital divide, human rights, social inclusion, economic sustainability and life-long learning. However, we can see that such indicators are very general and it is hard to incorporate them in a national context and determine their actual impact and correlation with e-government development. Despite economic activities on national level could significantly affect e-government development in individual country, research [11][41][42] dealing with national economic indicators and their implications for e-government is rather limited. Scarce studies [41] in the area are trying to define connections between national economic indicators and e-government development indicators specifically emphasizing correlations between GDP per capita, competitiveness, economic performance, government efficiency, use of ICT in the private sector, innovation index and internet access. The aspect of public interest within outlined methodologies is poorly elaborated and thus inadequately evaluated, while it appears that identification and formulation of vaguely indicated long-term public goals is hardly reached by public consensus.

4) Evaluation of e-government policies – issues and barriers

Evaluation of e-government policies is generally difficult [6][8][27][28], given the numerous obstacles to evaluation (Table 1) [43], complexity of public interest and frequent lack of clarity of objectives owing to the different and often competing stakeholder views. In addition, overlapping of initiatives and policies and their continuous fine-tuning related to volatile public opinion complicate monitoring and evaluation. The fact that e-government is relatively new is probably the main reason for fewer models and actual outcome experiences that can be used for benchmarking [43] and inclusion of public interest dimension. Effective evaluation including aspect of public interest requires good metrics, regular monitoring and reporting, disciplined and professional use of robust evaluation frameworks and the use of long-term evaluation practices. These qualities depend on a government's overall evaluation culture [43]. E-government project failures could have been mitigated by appropriate and comprehensive evaluation in the course of their conceptualization and planning [12][44][45], the identification and elimination of the main obstacles to e-government evaluation, which obviously extend to several areas, such as: institutional, political, social, and cultural area, will require a broad consensus and strong commitment of all stakeholders.

TABLE I. OBSTACLES TO E-GOVERNMENT EVALUATION

Obstacle	Example
Lack of clarity of objectives - stated goals may not have associated measures of progress; there may be multiple objectives	Hard to measure "quality of life".
Hard to define success	If people are spending more time online, is that good or bad?
Easy to be too ambitious	Several countries have set targets of "all services online" by specific dates. But not all services are appropriate to put online.
Information paradox	The benefits of ICT investment may not be visible for some time.
Question of who are the clients; multiple clients	Should one evaluate benefits for the users, the employees, the government at large, partners, etc.?
Hard to measure shared benefits	Shared infrastructure, multiple projects benefiting from shared portal, etc.
Private sector tools may not work for governments	Governments place importance on social values that are not incorporated into private sector tools and objectives.
Available indicators may not be the good ones	Current indicators (such as number of employees with internet connections) are helpful, but have limits.
Government definitions and methodologies vary from one country to the next	Collecting data is easier at the local level, but at that level administrations are highly decentralised.
Incentives to misstate evaluation results	If an organisation succeeds in saving money, telling others may result in their losing that money.
Challenge of sharing results	Hard to get organisations to report unsatisfactory results.
What you measure may become focus of organisation	If you measure number of services online, but not service quality, priority will be on putting services online but not on service quality.

III. ANALYSIS OF EXISTING METHODOLOGIES FOR EVALUATION OF E-GOVERNMENT POLICIES

The review of existing methodologies and various alternative approaches for evaluation of e-government policies was conducted in the second half of 2011. During that time we conducted an analysis of primary and secondary online resources, policy papers, reports, books, strategic documents, action plans and other documents containing e-government related research. In the initial phase of the review, we used key words of evaluation, assessment, measurement, monitoring, indicator models, e-government projects, e-government policies and effects (impacts) of e-government policies, and subsequently identified and

 ISBN: 978-1-61208-228-8

retrieved more than 380 related references in total. The frequency of references is becoming much higher in the second half of the last decade, proving the field is evolving rapidly and the interest of both the research community as well as policy makers on national and international level is increasing. In the second phase of the review, the identified references were tested by inclusion into the research framework containing two criteria, namely 1) identified reference must be completed project where evaluation of e-government policies is clearly outlined as the main research objective and 2) the reference must contain explicit indicators or benchmarks for evaluation of e-government policies. References which did not comply with both criteria as set out in the research framework were eliminated. After substantive verification and filtration, the vast majority of the items were excluded, leaving only 52 valid references. Identified methodologies and their publication types are catalogued in Table 2.

TABLE II. IDENTIFIED METHODOLOGIES FOR EVALUATION OF E-GOVERNMENT POLICIES

Author(s) – Year	Publication type
[30] [31] [32] [33] [46] [47] [48] [49]	Handbook / Tool
[9] [14] [24] [43] [50] [51] [52]	Policy paper
[11] [34] [35] [36] [38] [39] [40] [41] [42] [53] [54] [55] [56] [57] [58] [59] [60] [61] [62] [63] [64] [65] [66] [67] [68] [69] [70] [71] [72] [73]	Academic paper / Book
[13] [15] [74] [75] [76] [77] [78]	Report

Conducting our review we have identified basically three types of references dealing with our subject of discourse. Taking into account their development level we categorized the identified methodologies into three groups: 1) purely theoretical papers aiming to develop some kind of conceptual framework for evaluation of e-government policies, 2) research efforts developed up to the degree of pilot application and 3) methodologies developed in the practice for the practice (practical application).

Analyzing the diverse variety of evaluation methodologies identified in this area, certain general characteristics were identified and summarized below:

- Majority of the identified methodologies (30) for evaluation of e-government policies are presented in scholarly papers and books.
- Certain methodologies are rather abstract containing speculatively selected indicators often encompassing non evidence-based theoretical platforms, while their utilization does not facilitate the acquirement of quantifiable evaluation results.
- Accredited methodologies are to a large extent narrowly focused assessing predominantly one of the various evaluation aspects.
- More mature methodologies are consisted of a large number of indicators, which are normally aligned for evaluation of e-government policies in the originating countries.
- Majority of the identified methodologies are not providing a comprehensive evaluation of complex e-government policies impacts and their potential long-term outcomes.

IV. PRESENCE OF THE PUBLIC INTEREST ASPECT IN THE EXISTING METHODOLOGIES

Quest for satisfaction of the public interest, as presented in theory [16][22][23], should be at the heart of every e-government policy making process, while its various dimensions should be comprehensively covered in the setting of long-term public goals. The latter assumption requires accountability of public policy makers and evidence-based decision making, which must be based on comprehensive and balanced methodologies facilitating evaluation of various aspects of the designated e-government policy, its effects and the potential far-reaching consequences. Deliberation between public stakeholders [83], transparent policy making process and ultimate public consensus will increase the viability of high investments in e-government and facilitate positive response to the e-government policy and more beneficial acceptance of new e-services while allowing the pursuit of public interest and overall social development.

Proliferation of advanced ICT solutions and development of e-government have changed the social structure and political-sociological paradigm of the country as the widest social community [14][79]. E-government demonstrates considerable potential for the achievement of various ICT-induced social changes which could hopefully result in the accomplishment of at least some of the objectives within the public interest domain.

Political-sociological effects of ICT and e-government on the society in general are very complex. They have a significant impact on changes of the social environment, they are affecting old and creating new forms of work and changing perception of the world and social relations [72][80][81][82]. Accordingly, existing methodologies are converging on the following aspects of public interest in e-government evaluation: accessibility [4][24][55], citizens' trust and confidence [25][70][71], digital divide [4][43][15][79], social stratification and cohesion, citizens' rights and democratic participation [5][15][40], openness, transparency and corruption [9][14][24]. Notwithstanding the importance of adequate evaluation of e-government impacts [6][9], integrated methodologies covering the aspect of public interest comprehensively, are rather scarce. Findings regarding the public interest aspect in the existing methodologies are categorized below:

- The aspect of public interest is assigned a peripheral role in most of the existing methodologies for evaluation of e-government policies.
- The concept of public interest is not clearly elaborated and categorized in the existing methodologies, preventing its comprehensive inclusion in the actual evaluation undertakings.
- Methodologies are applying large number of indicators when focused on evaluation of the costs, benefits and risks (the most segments of public interest aspect are usually incorporated with

 ISBN: 978-1-61208-228-8

benefits). However, the aspect of public interest in its individual form is usually allocated a very small number of indicators.

- Covering public interest in existing methodologies is particularized, usually including arbitrarily selected dimension of public interest.
- Aspect of public interest presented in methodologies is often inadequate and superficial focusing predominantly on financial benefits and omitting all other societal aspects while reducing the potential of evaluation process for objective and comprehensive evaluation.
- Existing examples of the integration of public interest aspect in the methodologies are speculative, since the segments of the public interest within the methodology were developed by policy makers without appropriate deliberation procedure and public consensus.

Accordingly, inclusion of the public interest concept into the comprehensive methodology for evaluation of e-government policies should encompass the following activities:

- Analyze the overall evaluation field and define the particular aspects of evaluation within e-government policy (e.g., infrastructural aspect, organizational aspect, political aspect, etc.).
- Define and clearly structure the notion of public interest and associated components. Constructs should not be too abstract, because it could prevent the acquisition of the required data, establishment of indicators and their measurement in practice, decreasing the overall evaluation success.
- Concept of the public interest should constitute a relatively autonomous category, preventing the dissipation of its components between other aspects of evaluation, which could significantly diminish its importance.
- Delineate the aspects of evaluation as much as possible and prevent overlapping and transition of the indicators from different aspects of evaluation.
- Conceptualize adequate and measurable indicators containing precisely specified object and unit of measurement, structure, context, etc.
- Indicators should be specifically focused on the evaluation of long-term public interest and goals that have been set out in the designated e-government policy. Evaluation of the public interest concept should be multidimensional including the demand side of e-government services (user preferences, needs and satisfaction, etc.), general value of e-government policy for all social groups, its contribution to sustainable and inclusive social development, human rights and liberties, development of democratic values, etc.
- Assign appropriate weight to the concept of public interest and its components in the final aggregation of the evaluation results.

V. CONCLUSION AND FUTURE WORK

Research revealed a multitude of approaches which have been used in previous research studies and evaluation endeavors in this area, but still, their utilization in the process of comprehensive evaluation of e-government policies is only conditionally exploitable.

Despite limitations, seen mainly in the large diversity and absence of unified and clear theoretical framework, conducted analysis provides a valuable insight into the current e-government evaluation practice and facilitates exposure of inadequately evaluated public interest areas in the domain of e-government policies. The analysis results represent an advance in research of evaluation metrics and may eventually provide a solid platform for establishment of comprehensive methodology for evaluation of e-government policies including public interest aspect.

Problems in the development, usage and general success of e-government initiatives are evidently interrelated with the low quality and underdeveloped methodologies for evaluation of e-government policies and their public effects. Addressed shortcomings will have to be resolved, in order to ensure quality evaluation and disclosure of objective situation in the field, which could ultimately initiate the broader inclusion of public interest dimensions into the e-government policy-making procedures, and accelerate the overall development of e-government policies and appropriate e-services with added value for all stakeholders.

REFERENCES

[1] R. Silcock, What is E-government? “Parliamentary Affairs,” Hansard Society, vol. 54, pp. 88–101, 2001.

[2] OECD. Organization for Economic Co-operation and Development, E-government: Analysis Framework and Methodology, 2001.

[3] M. Yildiz, “Examining the Motivations for E-Government from an Institutional Theory Perspective: Evidence from Turkey,” The National Conference on Digital Government Research, Boston, 2003.

[4] OECD. Organization for Economic Co-operation and Development, Focus on Citizens: Public Engagement for Better Policy and Services. Paris, 2009.

[5] OECD. Organization for Economic Co-operation and Development, Rethinking e-Government Services: User-centred approaches. Paris, 2009.

[6] OECD. Organization for Economic Co-operation and Development, E-Government Indicators: Proposal for a new Framework and Methodology, Paris, 2010.

[7] Capgemini, Smarter, Faster, Better eGovernment, 8th Benchmark Measurement, European Commission, 2009.

[8] Capgemini, Rand, DTi, IDC, Sogeti, Digitizing Public Services in Europe: Putting ambition into action. Report of the 9th Measurement. European Commission, 2010.

[9] Capgemini, IDC, Rand Europe, Sogeti and DTi, eGovernment Benchmark Pilot on Open Government and Transparency. European Commission, 2011.

[10] M. Kunstelj, T. Jukić, and M. Vintar, “Analysing the demand side of e-government: what can we learn from Slovenian

 ISBN: 978-1-61208-228-8

users?" [EGOV 2007, LNCS 4656 (pp. 305–317). Wimmer, M., A. et al. (eds.), 2007].

[11] C. K. Kim, "A Cross-national Analysis of Global E-government," Public Organization Review, vol. 7, pp. 317–329, 2007.

[12] M. Vintar and J. Nograšek, "How much can we trust different e-government surveys? The case of Slovenia," Information polity, vol. 15, pp. 199–213, 2010.

[13] S. Dutta, "Estonia: A Sustainable Success in Networked Readiness?" The Global Information Technology Report 2006–2007, Palgrave Macmillan, pp. 81–90, 2007.

[14] UN, United Nations E-Government Survey, Leveraging e-government at a time of financial and economic crisis. Department of Economic and Social Affairs, 2010.

[15] WEF. World Economic Forum, The Global Information Technology Report 2010–2011. Transformations 2.0, 2011.

[16] C. Waelde and M. McGinley, "Public Domain; Public Interest; Public Funding: focussing on the 'three Ps' in scientific research," SCRIPTed, vol. 71, pp. 83–106, 2005.

[17] R. Yin, Case study research: design and methods (3rd edn.). Sage Publications, Thousand Oaks, USA, 2003.

[18] M. Patton, Qualitative Evaluation and Research Methods (2nd edn.). Sage Publications, Thousand Oaks, USA, 1990.

[19] G. Laurie and C. Waelde, "Privacy, Property and Personalities: Whatever Happened to the Public Interest?" Law & Contemp. Probs. 33, pp. 45–63, 2005.

[20] D. D. Nimmo and T. D. Ungs, American political patterns conflict and consensus. Boston, Little, Brown, 2nd ed., 1969.

[21] B. Ringer, "Authors' Rights in the Electronic Age: Beyond the Copyright Act of 1976," Loyola of Los Angeles Entertainment Law Journal 1 (1), 1981.

[22] R. E. Flathman, The Public Interest: An Essay Concerning the Normative Discourse of Politics, Wiley, New York, 1966.

[23] P. D. Gordon, Public Administration in the Public Interest. A prescriptive analysis of a democratic humanist paradigm of public administration, 1975.

[24] UN, United Nations E-Government Survey, From E-Government to Connected Governance. Department of Economic and Social Affairs, New York, 2008.

[25] Accenture, From e-Government to e-Governance: Using new technologies to strengthen relationships with citizens. Institute for Health and Public Service Value, Dublin, 2009.

[26] D. M. West, Improving Technology Utilization in Electronic Government around the World, Center for Public Policy, Brown University, USA, 2008.

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

[28] L. Dadayan, "Measuring Return on Government IT Investments," Proceedings of the 13th European Conference on Information Technology Evaluation. Genoa, Italy, 2006.

[29] Economist Intelligence Unit, Digital economy rankings 2010. Beyond e-readiness, 2010.

[30] ADAE, 2007, MAREVA methodology guide: Analysis of the value of ADELE projects. Fourth High Level Seminar on Measuring and Evaluating E-Government, Dubai, 2007.

[31] RSO SpA & LUISS Management, eGEP - eGovernment Economics Project: Expenditure Study Final Version. European Commission, 2006.

[32] P. Rothig. WiBe 4.0 Methodology. Economic Efficiency Assessments in Particular with Regard to the Use of Information & Communication Technology, 2004, 2010.

[33] AGIMO – Australian Government Information Office, 2004, Demand and Value Assessment Methodology - DAM & VAM. Commonwealth of Australia, 2004.

[34] B. Klievink and M. Janssen, "Realizing joined-up government – Dynamic capabilities and stage models for transformation," Government Information Quarterly, vol. 6, pp. 275–284, 2009.

[35] A. Fleur Van Veenstra, M. Janssen, and Y. H. Tan, "Towards an Understanding of E-Government Induced Change - Drawing on Organization and Structuration Theories," [EGOV 2010, LNCS 6228 (pp. 1–12). Wimmer, M. A. et al. (eds.), 2010].

[36] K. Schedler and B. Schmidt, "Managing the e-government organization," International Public Management Review, vol. 5, pp. 1–20, 2004.

[37] H. J. Scholl, "E-government: A Special Case of ICT-enabled Business Process Change," [36th Hawaii International Conference on System Sciences, 1/12/ 2003].

[38] H. J. Scholl, "Electronic Government: Information Management Capacity, Organizational Capabilities, and the Sourcing Mix," Government Information Quarterly, vol. 23, pp. 73–96, 2006.

[39] C. Leitner and M. Kreuzeder, "Organisational Changes, Skills and the Role of Leadership Required by eGovernment," [EGOV 2005, LNCS 3591, Wimmer, M. A. et al. (eds.), 2005].

[40] B. Martin and J. Byrne, Implementing e-Government: widening the lens. RMIT University, Melbourne, 2003.

[41] C. Bavec and M. Vintar, "What Matters in the Development of the E-Government in the EU?" [EGOV 2007, LNCS 4656, (pp. 424–435). Wimmer, M. A. et al. (eds.), 2007].

[42] H. Singh, A. Das, and D. Joseph, "Country-Level Determinants of E-Government Maturity," Communications of the Association for Information Systems, vol. 20, 2007.

[43] OECD. Working Group 2 on E-government and Administrative Simplification, Good Governance for Development in Arab Countries, Dubai, 2007.

[44] D. Janssen, S. Rotthier, and K. Snijkers, "If you measure it, they will score: An assessment of international eGovernment benchmarking," Information Polity, vol. 9, pp. 121–130, 2004.

[45] B. A. Hamilton, Building a Methodology for Measuring the Value of E-Services. USA Social Security Administration, 2002.

[46] CIO Council, Value Measuring Methodology: How-To-Guide. Washington: Best Practices Committee, 2002.

[47] OECD. Directorate for science, technology and industry, Working party on indicators for the information society. Guide to measuring the information society. Paris, 2005.

[48] T. P. Rama Rao, V. Rao, V. Bhatnagar, S. C. Satyanarayana, and J. Shri, E-Governance Assessment Frameworks (EAF Version 2.0). Government of India, 2004.

 ISBN: 978-1-61208-228-8

[49] S. Baudu and S. Dzhumalieva, Value Assessment Tool - VAST. Licensed by the European Commission under EUPL. European Commission, 2010.

[50] UNESCO, E-government Tool Kit for Developing Countries. Asia Pacific Regional Bureau for Communication and Information, New Delhi, 2005.

[51] Republic of Korea, Ministry of Public Administration and Security, e-Government in Korea. OECD e-Leaders Meeting 2010, 11-12 Oct. 2010, Brussels, 2010.

[52] European Commission, Capgemini, & IDC, eGovernment Benchmarking in 2011, 2011.

[53] D. K. Mahalik, "Outsourcing in e-Governance: A Multi Criteria Decision Making Approach," Journal of Administration & Governance, vol. 5, pp. 24–35, 2010.

[54] P. R. Devadoss, S. L. Pan, and J. C. Huang, "Structurational analysis of e-government initiatives: a case study of SCO," Decision Support Systems, vol. 34, pp. 253–269, 2002.

[55] A. T. K. Ho, "Reinventing Local Governments and the E-Government Initiative," Public Administration Review, vol. 62, pp. 434–444, 2002.

[56] M. J. Moon, "The Evolution of E-Government among Municipalities: Rhetoric or Reality?" Public Administration Review, vol. 62, pp. 424–433, 2002.

[57] Y. C. Chen and J. Perry, "Outsourcing for E-Government: Managing for Success," Public Performance & Management Review, vol. 26, pp. 404–421, 2003.

[58] O. O'Donnell, R. Boyle, and V. Timonen, "Transformational aspects of E-government in Ireland: Issues to be addressed," Electronic Journal of e-Government, pp. 23–32, 2003.

[59] D. Griffin, A. Foster, and E. Halpin, "Joined-up E-government: an exploratory study of UK local government progress," Journal of Information Science and Technology, pp. 58–83, 2004.

[60] M. Elnaghi, S. Alshawi, and F. Missi, "A Leadership Model for e-Government Transformation," [European and Mediterranean Conference on Information Systems, 2007].

[61] M. I. Štemberger and J. Jaklič, "Towards E-government by business process change – A methodology for public sector," Int. J. Inf. Manage., vol. 27, pp. 221–232, 2007.

[62] C. Pollitt, "Technological Change: a central yet neglected feature of public administration," [NISPA10, 2010].

[63] K. M. Boyer-Wright and J. E. Kottemann, "An Empirical Assessment of Common Fundamentals in National E-Readiness Frameworks," Journal of Global Information Technology Management, vol. 12, pp. 55–74, 2009.

[64] L. Picci, "The quantitative evaluation of the Economic Impact of E-Government: A structural Modelling Approach," Information Economics and Policy, vol. 18, pp. 107–123, 2006.

[65] D. Y. Rhee, "The Study on Determinants of E-Government Evolution," Journal of Korean Association for Regional Information Society, vol. 12, pp. 183–201, 2009.

[66] R. W. Schwester, "Socio-Demographic Determinants of E-government Adoption: An Examination of Major U.S. Cities," Journal of Public Management and Social Policy, vol. 16, pp. 21–32, 2010.

[67] C. Serrano-Cinca, M. Rueda-Tomás, and P. Portillo-Tarragona, "Determinants of E-Government Extension," Online Information Review, vol. 33, pp. 476–498, 2009.

[68] I. Gallego-Álvarez, L. Rodríguez-Domínguez, and I. M. García-Sánchez, "Are Determining Factors of Municipal E-Government Common to a Worldwide Municipal View? An Intra-Country Comparison," Government Information Quarterly, vol. 27, pp. 423–430, 2010.

[69] K. Layne and J. Lee, "Developing fully functional E-government: A four stage model," Government Information Quarterly, pp. 122–136, 2001.

[70] C. J. Tolbert and K. Mossberger, "The Effects of E-Government on Trust and Confidence in Government," Public Administration Review, vol. 66, pp. 354–70, 2006.

[71] M. Parent, C. A. Vandebeek, and A. C. Gemino, "Building Citizen Trust through e-Government," Government Information Quarterly, vol. 22, pp. 720–736, 2005.

[72] D. C. Shim and T. H. Eom, "E-Government and Anti-Corruption: Empirical Analysis of International Data," International Journal of Public Administration, vol. 31, pp. 298–316, 2008.

[73] S. Bolgherini, "The Technology Trap and the Role of Political and Cultural Variables: A Critical Analysis of the E-Government Policies," Review of Policy Research, vol. 24, pp. 259–275, 2007.

[74] Deloitte & Ipsos Belgium, Benchmarking deployment of eHealth services III, 2011.

[75] P. Wauters, M. Nijskens, and J. Tiebout, Capgemini, The User Challange, Benchmarking The Supply Of Online Public Services, 2007.

[76] I. Meyer, T. Hüsing, M. Didero and W. B. Korte, (Empirica), eHealth Benchmarking (Phase II), 2009.

[77] R. Weehuizen and C. van Oranje, Innovative and adaptive pan-European services for citizens in 2010 and beyond Pan-European eGovernment Services (PEGS) in perspective: function, forms, actors, areas, pathways and indicators. European Commission, 2007.

[78] M. Holzer and S. T. Kim, Digital Governance in Municipalities Worldwide, A Longitudinal Assessment of Municipal Websites Throughout the World, The e-governance institute, New Jersey, 2007.

[79] P. Norris, Digital Divide: Civic Engagement, Information Poverty and the Internet, Cambridge University Press, Cambridge, 2001.

[80] J. Slevin, The Internet and Society, Polity Press, Cambridge, 2000.

[81] A. V. Anttiroiko, "Toward the European Information Society," Communications of the ACM, vol. 44, pp. 31–36, 2001.

[82] C. Thurlow, L. Lengel, and A. Tomic, Computer mediated communication: Social interaction and the internet. London, Thousand Oaks, New Delhi, Sage Publications, 2004.

[83] P. J. Lor and J. J. Britz, "Is a knowledge society possible without freedom of access to information?" Journal of Information Science, vol. 33, pp. 387–397, 2007.

 ISBN: 978-1-61208-228-8

Evaluation set up for a Regional Social Network Service

Assessing content, social networks and user attitudes of the Austrian SNS "Sankt Onlein"

Isabel Anger, Thomas Rößler, Christian Kittl
evolaris next level GmbH
Graz, Austria
thomas.roessler@evolaris.net, christian.kittl@evolaris.net, isabel.anger@evolaris.net

Abstract— **Social Network Services (SNSs) are vastly popular and used by millions of people worldwide. A few years ago, research has begun to analyze these platforms as well as the social networks and social interactions that are enabled by these technologies. Many of the well-known and popular SNSs such as Facebook or YouTube have already been analyzed and profiled under numerous different aspects. However, regional SNSs have largely been disregarded. In this work in progress paper, we introduce a newly launched regional SNS from Austria, outline the particularities of this regional social networking platform, set up first hypotheses and outline our evaluation approaches to analyzing this service. It is our believe that by putting the particularities of a regional SNS in focus, and by using an interaction graph that visualizes intimate social links instead of the commonly used classical social network analysis, valuable insights on the connection of online networking activities and offline social interactions as well as on the users' perception of familiarity and place can be obtained. The paper illustrates the particularities of regional SNS and describes an evaluation set up to gain more insights into this specific type of networking platform.**

Keywords - Social Network Service; social network analysis; content analysis; evaluation; measurement.

I. INTRODUCTION

Close to 900 million people actively use Facebook every month. The need for using digital SNSs to share content and interact with other people seems to be consistent. Facebook is, however, struggling with extensive privacy and data mining discussions and also globalizes social interaction possibilities, making it thus more complex to connect online social networks with offline social networks. This fuels the development of alternative, local SNSs that encompass country- or region-specific differences in their functionalities and the language used in the user interface. An example of a very successful and well-established regional SNS is Qzone, a platform developed for Chinese users with an estimated number of 480 million users of whom 150 million users update their accounts at least once a month [1]. Qzone provides blogging, media sharing, and social networking functions additionally to an instant messenger which is deemed to be one of the key success factors of the platform. Another factor of success is deemed to be the fact that Qzone is reserved for Chinese users, supporting the Chinese language only. Thus, this regional SNS can hardly be compared with Facebook: Although the platforms partially offer the same functionalities to their users, the boundaries of the two SNSs differ in terms of language and cognitive confinements. Whereas Facebook is open to any Internet user above the age of 13 with a valid email address, supporting a broad selection of languages and users from all over the world, Qzone demands first and foremost the knowledge of the Chinese language of the user.

This concept seems to be a factor of success: As reported by analysts and industry professionals, Qzone is more important than Facebook in China. Apart from the example Qzone, other regional-specific SNSs have established themselves in confined regions all over the world, among them vkontakte in Russia, Odnoklassniki in the Ukraine and Kazakhstan, and Mixi in Japan [2].

This leads to our research questions: Is a local, language-specific or regional-specific SNS more appealing to users? Does a regional SNS enable a deeper connection of their online social network with offline social activities? Can a regional SNS be characterized by more intimate node connections and thus more interaction-intensive social networks?

In order to verify our assumptions, we plan to conduct several evaluations which will be outlined in this work in progress paper. At first, the functionalities and characteristics of a regional SNS will be described in more depth; then, state-of-the-art evaluation methods for social networks will be presented, and in the third part, the planned evaluation set up will be explained. Finally, we provide an outlook on further research topics connected to this topic.

II. SANKT ONLEIN: A REGIONAL SNS

In order to verify our hypotheses, we chose to analyze Sankt Onlein, a SNS which launched in 2011 and can be described as a profile-centric, identity-driven SNS [3]. Just as the very successful Qzone, Sankt Onlein is set up for a specific target group, which is comprised of the Austrian population. This limitation is, however, not forced by regulatory restrictions but instead promoted with the

 ISBN: 978-1-61208-228-8

proposition as being the "digital capital of Austria". This is underlined by signature terms from typical Austrian vocabulary being used for different functionalities and descriptions. The Austrian idioms are dialect forms of standard German, belonging to Bavarian family of languages for the most part.

Sankt Onlein also tries to diversify itself from other networks by providing a comprehensive set of privacy and data protection settings to its users, thus setting up a more trustful environment for social interaction. For example, Sankt Onlein promises to erase all user data and interaction records whenever a user chooses to delete his or her account. Additionally, Sankt Onlein strives to draw a greater value from user-generated content. With "Der Onleiner", Sankt Onlein has created a user-generated online newspaper that draws its contents from the platform's members. Sankt Onlein is currently used by more than 10.000 users situated in Austria. Mobile applications for smartphones are also offered: They provide the same functionalities as the website and are available for iOS devices and Android phones.

Users on Sankt Onlein are able to set up a profile for themselves, share media such as photos, add other platform members to their contacts, and interact with these contacts in a private or semi-public manner. Public posting, i.e., so that individuals who are not registered on the platform can view a message, or seeing a user's profile without registration is not possible. A user can choose to categorize his contacts into three privacy circles. This enables him to choose which circle a status message or shared medium is addressed to. Interaction possibilities include semi-public messages on the user's stream as well as comments on posted semi-public messages and published photos – this content appears on the user's "Schaukasten" a stream of activities on the user's profile page. Additionally, it is possible to send private messages, and finally, special, non-written messages in form of interaction buttons that allow a user to "wink at", "pad shoulder of", and "flirt with" another user.

Additionally to these user-to-user interactions, Sankt Onlein also provides a tool for creating community pages for companies, clubs, and other organizations. Users can express their sympathies for such a page by clicking the corresponding button and interacting with the community members afterwards.

III. THEORETICAL BASIS AND THEORETICAL BASIS

At the present, SNSs are beginning to be of rising interest to the research community. Various studies have profiled different platforms such as YouTube, Flickr, and Orkut [5] and used classical social network analysis methods for gaining deep insights. In this approach, individuals are nodes in a network where links between nodes are formed through the state of being a "contact", "friend", or "buddy". One of the most comprehensive social graphs is the visualization of data from Microsoft Messenger which shows 180 million nodes and 1.3 billion undirected edges [6].

For the evaluation of Sankt Onlein, we have chosen a slightly different approach that also encompasses other methods of evaluation and therefore broadened the theoretical basis as follows.

The characteristics described in the prior chapter (regional focus, use of dialect forms, and self-positioning of the platform) imply that Sankt Onlein and the interests of its users are focused on Austrian users, culture, and news. This leads to our first hypothesis:

H1: Members of the SNS Sankt Onlein primarily discuss topics that focus on Austrian users, culture, and news.

Recent studies [4] have shown that in large SNSs such as Facebook, most users only interact with about 50 percent of their contacts (so-called "friends", however, this term does not infer to a real friendship underlying the connection). 20 percent of a Facebook user's friends are responsible for 70 percent of all interactions on the platform. This observation can be applied to about 90 percent of all Facebook users. This corroborates the findings that networks on Facebook are very widespread across regional, linguistic and interest fields. A regional SNS such as Sankt Onlein on the other hand, has by our definition more defined boundaries, limiting possible social actors to a regional and linguistic group. We assume that this prohibits a direct comparison of Facebook and Sankt Onlein and thus formulate our second hypothesis:

H2: In local SNS, social networks and social interactions shift from a global level to a more regional level, resulting in a higher number of and more intimate node connections. A regional SNS acts primarily as a tool and enabler for interactions in the offline sphere.

Digital social networking platforms or communities may be classified in two main categories in terms of the structure of social networks on the respective platform [3]. On the one hand, there are networks focused on facilitating novel acquaintances through shared interests or shared objectives, e.g., dating communities or passion-driven special interest communities. On the other hand, there are networks that enable transferring existing social networks and node connections onto a digital platform. Profile-centric platforms that are not focused on special interests fall under this category. Sankt Onlein can also be positioned here, although it does indicate some functions (e.g., "wink at", "flirt with") used for making first contact with possible new acquaintances. Sankt Onlein as a local, profile-centric SNS with more intimate node connections thus has another implication which we summarize under the following hypothesis:

H3: Consequently to H1, local SNS facilitate a better connection of online networks and activities; and offline networks and activities.

In their theory of virtual space and place [7], the authors show how the perception of space and place inside virtual worlds influence actions and enjoyment of the users. Continuing our assumption of Sankt Onlein to be a local SNS with the above-mentioned cognitive limitations, it can be positioned as a confined areal within the web. The platform's own slogan "Austria's digital capital" also indicates a certain positioning as a defined space. The third indicator for Sankt Onlein to be a confined place rather than

 ISBN: 978-1-61208-228-8

an open space as per definition in [7] is the fact that Sankt Onlein does not offer platform-enabled functionalities outside of its website. While large services such as Google or Facebook are ubiquitous in the web, enabling registration, logging in, commenting and sharing functionalities on third-party platforms, Sankt Onlein users can only interact and use the platform's functionalities while using the SNS itself. This eventually creates a homely atmosphere of familiarity [8]. Thus, our third hypothesis is as follows:

H4: Sankt Onlein is a confined areal within virtual space and can be considered a place which eventually evokes an atmosphere of familiarity.

IV. EVALUATION SET UP

For evaluating and verifying our theses, our evaluation set up contains three different aspects that reach from content analysis to a quantitative and qualitative social network analysis. In detail, our evaluation set up is as follows:

A. Part 1: Content Analysis

In order to verify our first hypothesis that Sankt Onlein is indeed focused on Austrian news, culture and content, and thus qualifies for a region-focused SNS, a content analysis of shared messages will be conducted. The status messages posted on the users' "Schaukasten" will be sampled, unitized, and coded manually. The decision to complete the content analysis manually instead of using automated text mining and entity extraction is based on the fact that automatic content analysis for German texts is not yet sophisticated enough to produce viable results – especially in case of Austria-specific idioms. In the final analysis, the data are assessed quantitatively and qualitatively. As a result of the analysis, we hope to find whether there is a significant tendency for Sankt Onlein users to converse about Austria-specific topics. Through the qualitative analysis of the sample we also hope to see first indicators of whether the dialect currently used on the platform to describe functionalities and trigger actions is the same dialect as used in the user-produced content.

B. Social Network Analysis

In order to verify our second hypothesis that a regional SNS facilitates social networks with a higher number of and more intimate nodes, a social network analysis will be conducted. Not only will the amount of social links be analyzed, but also the quality of the links measured in user interactions. For this purpose, we plan to base our study on the concept of interaction graphs as presented in [4]. The expected results will provide insights on the network structure on Sankt Onlein. These results may be compared to data extracted from Facebook to underline the differences between the two SNSs. An especially interesting topic in this context is the comparison of an individual's network on different platforms, e.g., on Facebook and on Sankt Onlein, to find usage patterns and first qualitative insights into discussed topics.

C. Online User Survey

In the last part of the evaluation set up, an online user survey will be conducted. Through questioning users of Sankt Onlein about their usage patterns of the platform and their attitude towards this SNS, we hope to gain more qualitative insights and verify our theses 3 and 4. For designing the survey, we will build upon the use of the expectation disconfirmation theory (EDT) as presented by in [9] in the context of IT systems. In this approach, cognition and behavior is taken into account as well as temporal changes in technology acceptance models. We plan to survey two different user groups: One group G_1 which is already familiar with the platform and has been using it for some time; and one group G_2 which is new to the platform. The questionnaires will be sent to the users at two different points in time t_1 and t_2, and will survey perceptions, attitudes, and satisfaction of the users. The questionnaire for G_2 at time t_1 assesses the perceptions and attitude towards the SNS prior to use. The data gained in this first survey will be compared to the information gained from the questionnaire at t_2, in which disconfirmation, satisfaction and modified perceptions and attitudes as well as behavioral changes are surveyed.

V. CONCLUSION AND FUTURE WORK

In this short paper, the regional SNS Sankt Onlein was introduced and preliminary hypotheses on the usage characteristics of this platform were proposed. We presented our planned research activities concerning this regional SNS in hope of gaining valuable insights for learning more about this special type of SNS, as we have realized in our research and literature review activities, most research work focuses on large social network such as Facebook. A focus on smaller, regionally bordered SNSs may help to delve deeper into psychological, socio-scientific, and interaction issues that arise in the crossover zone between online social interaction and offline social interactions. These insights could be helpful for researchers and practitioners seeking to use regional SNS for e-commerce activities alike.

However, our approach represents only a small aspect of the many different research aspects of this SNS and can be seen as the first step towards deeper investigations. Comparison studies of other SNSs as mentioned in the introduction (vkontakte, Odnoklassniki, Mixi) could provide insights on generalization of regional SNS characteristics. However, possible limitations caused by data availability must be dealt with.

Also, individual aspects could be assessed in more detail. For example, linguistic research within the platform could result in first indicators about changes, shifts, and/or unifications of Austria-specific idioms.

Future work could also include a view on the mobile application of Sankt Onlein. Mobile penetration in Austria currently amounts to about 150 percent and mobile usage of SNS is constantly rising (as can be observed with Facebook and mobile-specific networks, e.g., Badoo). These statistics could suggest that the connection of online and offline social networks and interactions may be facilitated by the mobile

 ISBN: 978-1-61208-228-8

phone and SNS applications. As the mobile phone is "always on", it creates a state of ubiquitous social networking activities.

REFERENCES

[1] N. N., "The others: The world's other social media sites", The Financialist, 14 June 2012, retrieved from http://www.thefinancialist.com/the-others-the-worlds-other-social-media-sites/.

[2] V. Cosenza, "World Map of Social Networks", VincosBlog, 11 June 2012, retrieved from http://vincos.it/world-map-of-social-networks/.

[3] D. Boyd and N. Ellison, "Social Network Sites: Definition, History, and Scholarship", Journal of Computer-Mediated Communication Vol. 13, 2008, pp. 210-230, doi:10.1111/j.1083-6101.2007.00393.x.

[4] C. Wilson, B. Boe, A. Sala, K. Puttaswamy, and B. Zhao, "User Interactions in Social Networks and Their Implications", Proc. ACM European Conference on Computer Systems (EuroSys 09), ACM New York, 2009, pp. 205-218, doi:10.1145/1519065.1519089.

[5] A. Mislove, M. Marcon, K. Gummadi, P. Druschel, and B. Bhattacherjee, "Measurement and Analysis of Online Social Networks", Proc. Internet Measurement Conference (IMC 07), ACM New York, 2007, pp. 29-42, doi:10.1145/1298306.1298311.

[6] J. Leskovec and E. Horvitz, "Planetary-Scale Views on a Large Instant-Messaging Network", Proc. International World Wide Web Conference (WWW2008), ACM New York, 2008, doi: 10.1.1.133.6614

[7] C. Saunders, A. Rutkowski, M. van Genuchten, D. Vogel, and J. Molina Orrego, "Virtual Space and Place: Theory and Test", MIS Quarterly Vol. 35 Nr. 4, December 2011, pp. 1079-1098.

[8] D. Gefen and D. Straub, "Consumer Trust in B2C E-Commerce and the Importance of Social Presence: Experiments in E-Products and E-Services", Omega 32 Nr. 6, December 2004.

[9] A. Bhattacherjee and G. Premkumar, "Understanding Changes in Belief and Attitude Toward Information Technology Usage: A Theoretical Model and Longitudinal Test", MIS Quarterly, Vol. 28 Nr. 2, June 2004, pp. 229-254.

 ISBN: 978-1-61208-228-8

Social Privacy Protector - Protecting Users' Privacy in Social Networks

Michael Fire, Dima Kagan, Aviad Elishar, and Yuval Elovici
Telekom Innovation Laboratories and Information Systems Engineering Department,
Ben-Gurion University of the Negev, Beer-Sheva, Israel
Email: {mickyfi,kagandi,aviade, elovici}@bgu.ac.il

***Abstract*—In recent years, online social networks have grown exponentially, as these networks are fantastic places to meet and network with people who share similar personal interests. Facebook, currently the largest social network, has more than 901 million active users. The amount of personal information each user exposes on social networks like Facebook is staggering. Recent research in the area of social networking evaluated that many Facebook users exposed personal information. Due to the many security concerns regarding online personal exposure, we developed the Social Privacy Protector, a software which aims to improve the security and privacy of Facebook users. The software contains three protection layers which improve user privacy by implementing different methods. The software first identifies a user's friends who might pose a threat, and then restricts this "friend"'s exposure to the user's personal information. The second layer is an expansion of Facebook's basic privacy settings based on different types of social network usage profiles. The third layer alerts the user about the number of installed applications on their Facebook profile which have access to their private information. An initial version of the Social Privacy Protection software was evaluated on 74 Facebook users and successfully assisted them in restricting the access of 392 friends.**

Keywords-Social Network Analysis; Social Network Privacy; Social Network Security; Facebook Application; Fake Profiles.

I. INTRODUCTION

In recent years, online social networks have grown rapidly and today offer individuals endless possibilities for publicly expressing themselves, communicating with friends, and sharing information with people across the world. A recent survey [1] estimated that 65% of adult internet users use online social networks sites, such as Twitter [2], LinkedIn [3] Google+ [4], and Facebook [5]. As of June 2012, the Facebook social network, has more 955 million monthly active users [6]. On average, Facebook users have 138 friends and upload more than 300 million pictures each day [6]. Moreover, according to the Nielsen "Social Media Report" [7], American internet users spent more than 53.5 billion minutes on Facebook in the month of May 2011, making Facebook the leading web-brand in United-States.

Due to the friendly nature of Facebook, users tend to disclose many personal details about themselves and about their connections. These details can include date of birth, personal pictures, work place, email address, high school name, relationship statuses, and even phone numbers. Moreover, Bosmaf et al. [8] discovered that an average of 80% of studied Facebook users accepted friend requests from people they do not know if they share more than 11 mutual friends. In many cases, accepting friend request from strangers may result in exposure of a user's personal information to third parties. In addition, personal information of Facebook users can be exposed to third party Facebook applications [9]. Another privacy concern deals with existing privacy settings which, for the majority of Facebook users, do not match security expectations [10]. This results in many users accidently or unknowingly publishing private information, leaving them more exposed than they assumed.

If a user's personal information is disclosed to a malicious third party, it can be used to threaten the well-being of the user both online and in the real world. For example, a malicious user can use the gained personal information and send customized spam messages to the user in an attempt to lure such users onto malicious websites [11] or blackmail them into transferring money to the attacker's account [12]. In order to cover their tracks, social network attackers can use fake profiles. In fact, the number of fake profiles on Facebook can be counted in the tens of millions. Facebook estimates that around 5%-6% of its users could be false or duplicate accounts [13].

In this paper, we present an application for protecting user privacy on Facebook. Our application provides Facebook users with three different layers of protection. The first layer enables Facebook users an easy method for controlling their profile privacy settings by simply choosing the most suitable profile privacy settings in just one click. The second layer notifies the user of the number of applications installed on their profile that may impose a threat to their privacy. The third layer analyzes the user's friends list to identify which friends of the user are suspected as fake profiles and therefore impose a threat on a user's privacy. The application presents a convenient method for restricting the access of these fake profiles to the user's personal information without removing them from the user's friends list.

The remainder of this paper is organized as follows. In Section II, we give a brief overview of various related solutions which better help protect the security and privacy of social network users. In Section III, we describe the Social Privacy Protector software architecture in detail. In Section IV, we describe the initial evaluation results. Finally, in Section V, we present our conclusions from this study and

 ISBN: 978-1-61208-228-8

offer future research directions.

II. Related Work

In recent years, due to the increasing number of privacy and security threats on online social networks users, social network operators, security companies, and academic researchers have proposed various solutions to increase the security and privacy of social network users.

Social network operators attempt to better protect their users by adding authentication processes to ensure that the registered user represents a real live person [14]. Many social network operators, like Facebook, also offer their users a configurable user privacy setting that enables users to secure their personal data from other users in the network [10], [15]. Additional protection may include defense against hackers, spammers, socialbots, identity cloning, phising, and many other threats. For example, Facebook users have an option to report other users in the network who harass others in the network [16].

Many commercial and open source products, such as Checkpoint's SocialGuard [17], Websense's Defensio [18], UnitedParents [19], RecalimPrivacy [20], and PrivAware application [21], offer online social network users tools for better protecting themselves. For example, the Websense's Defensio software aims to protect its users from spammers, adults content, and malicious scripts on Facebook.

In recent years, several published academic studies have proposed solutions for different social network threats. DeBarr and Wechsler [22] used the graph centrality measure to identify spammers. Wang [23] presented techniques to classify spammers on Twitter based on content and graph features. Stringhini et al. [11] presented a solution for detecting spammers in social networks by using "honeyprofiles". Egele et al. [9] presented PoX, an extension for Facebook which make all requests for private data explicit to the user. Anwar and Fong [24] presented the Reflective Policy Assessment tool which helps the user examine their profile from the viewpoint of another user in the network. Recently, Fire et al. [25] proposed a method for detecting fake profiles in online social network based on anomalies in the fake user's social structure.

In this study, we offer a method for protecting user privacy in online social networks by detecting users who may pose a threat to a user's privacy by restricting their access to the user's personal information.

III. Social Privacy Protector Architecture

To better protect the privacy of Facebook users, we have developed the *Social Privacy Protector* software (otherwise referred to as SPP). The SPP software consists of three main parts (see Figure 1) which work in synergy: a) *Friends Analyzer Facebook application* - which is responsible for identifying a user's friends who may pose a threat to the users privacy, b) *SPP Firefox Addon* - which analyzes the user's privacy setting and helps the user to improve their privacy settings in one click, and c) *HTTP Server* - which is responsible for the analyzing, storing, and caching software results for each user. In the remainder of this section, we describe in detail each individual SPP part.

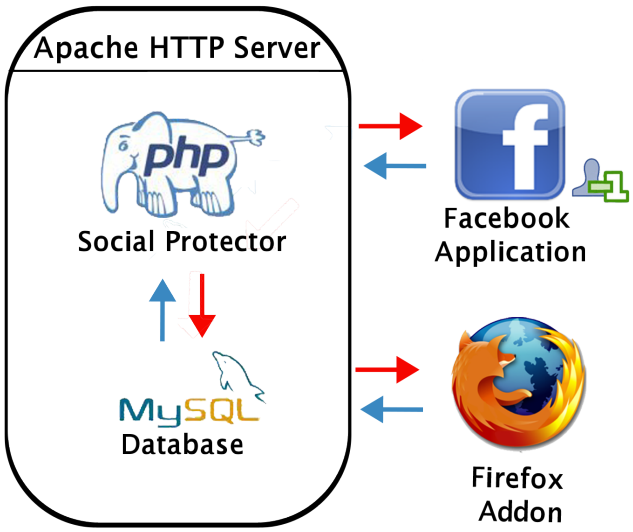

Figure 1. Social Privacy Protector Architecture

A. The Friends Analyzer Facebook Application

The Friends Analyzer Facebook application is the part of the SPP which is responsible for analyzing the user's friends list in order to determine which of the user's friends may pose a threat to the user's privacy. After the user installs the Friends Analyzer application, the application scans the user's friends list and returns a credibility score for each one of the user's friends. Each friend's score is created by simple heuristics which take into account the strength of the connection between the user and their friends. The application estimates the strength of each connection by calculating the number of common friends between the user and their friend, the number of pictures and videos the user and their friend were tagged in together, the number of groups the user and their friend were both members in, and the number of messages passed between the user and their friend. In the end of the process, the user receives a web page which includes a sorted list of all their friends. The list is sorted according to the score of each friend received where the friends with the lowest scores have the highest likelihood of being fake profiles appear on the top the list (see Figure 2). For each friend in the returned sorted list, the user has the ability to restrict the friend's access to the user's private information simply by clicking on the restrict button attached to each friend in the sorted list.

 ISBN: 978-1-61208-228-8

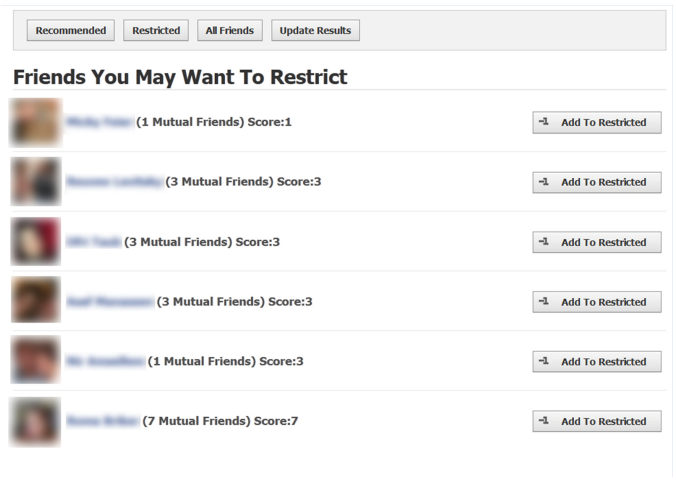

Figure 2. Friends Analyzer Facebook application - user's ranked and sorted friends list

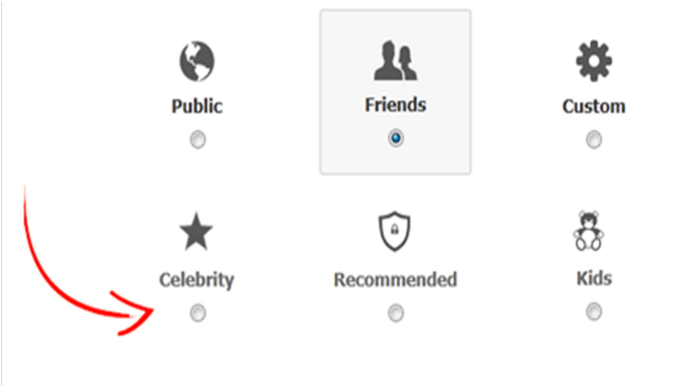

Figure 3. Social Privacy Protector Firefox Addon - optimizing the user's privacy setting in one simple click

B. Social Privacy Protector Firefox Addon

The Social Privacy Protector Firefox Addon (also referred to as Addon) is the part of the SPP software which is responsible for improving the user's privacy settings in just a few simple clicks. After the Addon is installed on the user's Firefox browser, it begins to monitor the user's internet activity. When the Addon identifies the user who has entered into their Facebook account, the Addon analyzes the number of applications installed on the user's Facebook profile and presents a warning with the number of installed applications that may pose threat to the user's privacy (see Figure 3).

The Addon also presents the top two results obtained by the Friends Analyzer Facebook application and suggests to the user which friends to restrict (see Figure 3).

The Addon also detects when the user has entered the Facebook's privacy settings page and presents the user with three new privacy setting options. The new privacy settings are based on the user's profile type and can be modified

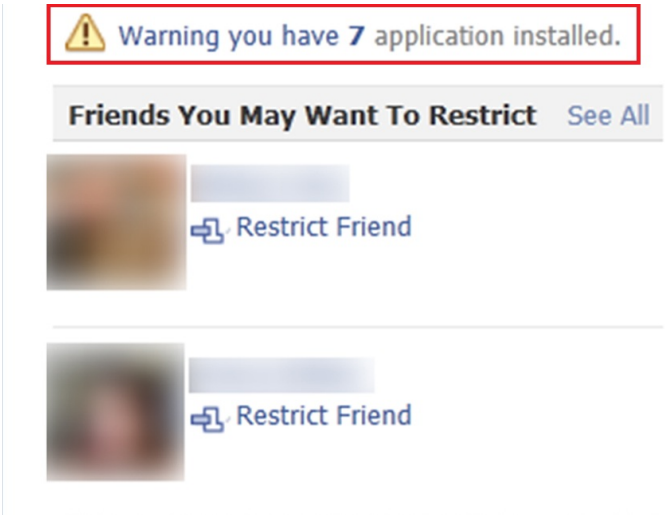

Figure 4. Social Privacy Protector Firefox Addon - warning about installed applications and friends you may want to restrict

in one click (see Figure 4) instead of the more complex Facebook custom privacy setting that may contain more than 170 options [24]. Using the new Addon privacy setting, a user can simply chose the profile type most suitable for him out of three options: a) *Celebrity setting* - in this setting all of the user's information is public, b) *Recommended setting* - in this setting the user's privacy is only public to friends, however some of the user's details, such as profile name and pictures, are public, and c) *Kids settings* - in this setting the profile is only open to the user's friends and only friends of friends can apply for friend requests. Using this Addon, users can easily control and improve their privacy without the need to contact a security expert. Our application is also easy for customizing privacy settings by adding more privacy option settings to different types of users.

C. HTTP Server

The HTTP Server is the part of the SPP which is responsible for connecting between the SPP Firefox Addon and SPP Facebook application. Moreover, to enhance the application's performance, the HTTP server caches parts of the analyzed results. In order to protect the user's privacy, the application stores only the minimal number of features in an encrypted manner using RC4 encryption.

IV. Evaluation

An initial version of the SPP version was evaluated by 74 users who installed the Friends Analyzer Facebook application and 4 users who installed the Addon. Using the Friends Analyzer Facebook application, 31 users have a restriction of 392 ($median = 3$ and $\sigma_{dev} = 25.76$) friends. According to our initial evaluation results, the average common friends between the user and the friends they chose to restrict was

 ISBN: 978-1-61208-228-8

12.82 and the average number of common tagged pictures was 0.14 (see Table I).

Table I
FRIENDS AND RESTRICTED FRIENDS STATISTICS

Feature	Restricted Friends	All Friends
Common-Friends Average	12.82	32.32
Common-Groups	0.36	0.684
Tagged Pictures	0.14	1.39
Common-Messages	1.31	3.14

V. CONCLUSIONS AND FUTURE WORK

In this paper, we present the SPP software which aims to better protect user's privacy in Facebook. The software protects user's privacy by providing three layers of privacy protection. The first layer helps to restrict a user's friends access to personal information. The second layer help to identify and warn the user about installed Facebook applications which can violate the user's privacy. The third layer helps the user to adjust their privacy setting in one click. An initial version of the software was evaluated by 74 users and helped 31 users restrict the access of 392 friends to the users' personal information. As expected, in our preliminary results (see Table I), the application users' chose to restrict friends with less common features with them, such as number of common friends and number of tagged pictures.

The study presented in this paper is a work in progress with many available future directions. Using the SPP we can gather examples on which friends users tend to restrict. Using these examples, combined with Machine-Learning algorithms, we can improve the recommendation results on which friends to restrict. Moreover, in case many users restricted the same users, we can conclude with high likelihood that these users are fake users and recommend Facebook to remove them from the social network. Another possible future direction is to collect anonymous data on user privacy setting preferences. By using the users privacy preferences, we can create more types of "one click" privacy settings that serve other types of users and protect their privacy.

In the near future, we are going to release a final stable version of the SPP, and make it available to any user that want to improve his privacy on Facebook.

VI. AVAILABILITY

The Social Privacy Protector and parts of its source code are available for download from http://www.socialprotector.net. The Friend Analyzer Facebook application is available to download from https://apps.facebook.com/friend_analyzer_app. A video with detailed explanations on how to use the SPP application is available in http://www.youtube.com/watch?v=Uf0LQsP4sSs

VII. ACKNOWLEDGEMENT

We would want to thank Jennifer Brill for repeated readings and markups on our grammar and spelling. We also want to thank the anonymous reviewers for their valuable comments and suggestions to improve the manuscript.

REFERENCES

[1] M. Madden and K. Zickuhr, "65% of online adults use social networking sites," http://pewinternet.org/Reports/2011/Social-Networking-Sites.aspx.

[2] Twitter, http://www.twitter.com/, [Online; accessed 1-September-2012].

[3] LinkedIn, http://www.linkedin.com/, [Online; accessed 1-September-2012].

[4] Google+, https://plus.google.com/, [Online; accessed 1-September-2012].

[5] Facebook, http://www.facebook.com/, [Online; accessed 1-September-2012].

[6] Facebook-Newsroom, http://www.facebook.com, [Online; accessed 1-September-2012].

[7] Nielsen, "The social media report," http://blog.nielsen.com/nielsenwire/social/, [Online; accessed 1-September-2012].

[8] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "The socialbot network: when bots socialize for fame and money," in *Proceedings of the 27th Annual Computer Security Applications Conference*. ACM, 2011, pp. 93–102.

[9] M. Egele, A. Moser, C. Kruegel, and E. Kirda, "Pox: Protecting users from malicious facebook applications," in *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2011 IEEE International Conference on*. IEEE, 2011, pp. 288–294.

[10] Y. Liu, K. Gummadi, B. Krishnamurthy, and A. Mislove, "Analyzing facebook privacy settings: User expectations vs. reality," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, 2011, pp. 61–70.

[11] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in *Proceedings of the 26th Annual Computer Security Applications Conference*. ACM, 2010, pp. 1–9.

[12] S. Nelson, J. Simek, and J. Foltin, "The legal implications of social networking," *Regent UL Rev.*, vol. 22, pp. 1–481, 2009.

[13] Facebook, http://www.sec.gov/Archives/edgar/data/1326801/000119312512101422/d287954ds1a.htm#toc287954_2.

[14] J. Kuzma, "Account creation security of social network sites," *International Journal of Applied Science and Technology*, vol. 1, no. 3, pp. 8–13, 2011.

 ISBN: 978-1-61208-228-8

[15] S. Mahmood and Y. Desmedt, "Poster: preliminary analysis of google+'s privacy," in *Proceedings of the 18th ACM conference on Computer and communications security*. ACM, 2011, pp. 809–812.

[16] Facebook, "Report abuse or policy violations."

[17] ZoneAlarm, http://www.zonealarm.com/, [Online; accessed 1-September-2012].

[18] W. Defensio, http://www.defensio.com/, [Online; accessed 1-September-2012].

[19] UnitedParents, http://www.unitedparents.com/, [Online; accessed 1-September-2012].

[20] ReclaimPrivacy, http://http://www.reclaimprivacy.org//, [Online; accessed 1-September-2012].

[21] PrivAware, http://apps.facebook.com/privaware/, [Online; accessed 1-September-2012].

[22] D. DeBarr and H. Wechsler, "Using social network analysis for spam detection," *Advances in Social Computing*, pp. 62–69, 2010.

[23] A. Wang, "Don't follow me: Spam detection in twitter," in *Security and Cryptography (SECRYPT), Proceedings of the 2010 International Conference on*. IEEE, 2010, pp. 1–10.

[24] M. Anwar and P. Fong, "A visualization tool for evaluating access control policies in facebook-style social network systems," 2012.

[25] M. Fire, G. Katz, and Y. Elovici, "Strangers intrusion detection - detecting spammers and fake profiles in social networks based on topology anomalies," 2012.

 ISBN: 978-1-61208-228-8

Complex Systems Engineering for Rapid Computational Socio-Cultural Network Analysis and Decision Support Systems

Tareq Z. Ahram
Institute for Advanced Systems Engineering (IASE)
University of Central Florida
Orlando, FL 32816, USA
tahram@ucf.edu

Waldemar Karwowski
Department of Industrial Engineering and Management Systems (IEMS)
University of Central Florida
Orlando, FL 32816, USA
wkar@ucf.edu

***Abstract*—The advent and adoption of internet-based social networking has significantly altered our daily lives. The educational community has taken notice of the positive aspects of social networking such as creation of blogs and to support groups of system designers going through the same challenges and difficulties. This paper introduces a social networking framework for collaborative education, design and modeling of the next generation of smarter products and services. Human behavior modeling in social networking application aims to ensure that human considerations for learners and designers have a prominent place in the integrated design and development of sustainable, smarter products throughout the total system lifecycle. Social networks blend self-directed learning and prescribed, existing information. The self-directed element creates interest within a learner and the ability to access existing information facilitates its transfer, and eventual retention of knowledge acquired. In conclusion, this research paper introduces the application of social-networking to design and modeling of products and services and provides a novel technology for facilitating the understanding of complex human behavior and to better identify crucial user needs.**

Keywords - social modeling; decision support systems; behavioral modeling.

I. INTRODUCTION

The concept of smartness of consumer products and services has been investigated by several authors. This section presents a synthesis and summary of the most innovative work that influenced research in this field. Allmendinger and Lombreglia [1] highlighted smartness in a product from a business perspective. They regard "smartness" as the product's capability to predict business errors and faults, thus "removing unpleasant surprises from [the users'] lives." Ambient Intelligence (AMI) group [2] describes a vision where distributed services, mobile computing, or embedded devices in almost any type of environment (e.g., homes, offices, cars), all integrate seamlessly with one another using information and intelligence to enhance user experiences [3, 4, 5, 23]. The advent and adoption of internet-based social networking sites such as MySpace TM and Facebook TM has significantly altered social interactions of their users. Users of social networking sites vary their activities; some may be very active sharing their daily life experiences with comments and pictures, while others simply use the sites as a personal directory service. The educational community has taken notice of the following positive aspects of social networking:

- Peer feedback, increasingly fast response times for scientific discovery, collaborative design and research
- Creation of blogs and support groups of individuals going through the same or similar difficulties
- Providing a social context in line with the university, company, design group, or field of study
  A venue with links not directly related to a given educational alignment or resource.

Social networking applications support the development of a methodology to better assess and predict imprecision and variability in user behavior by applying advanced mathematical and soft computing techniques to aid in studying human social, cultural and behavioral aspects. Application of soft computing techniques helps identify erroneous, problematic activities and issues that might otherwise go undetected for their obscurity, complexity, or elaborate inter-relationships. In addition to these above, social networks themselves are highly adaptable, flexible, and mobile. For example, the "blogging" paradigm became the "micro-blogging" concept known as Twitter, which now is integrated with Facebook's "status updates." Arguably, social networking provides an effective method of satisfying the primal human desire of communication.

Rapid technological advancements and agile manufacturing created what is called today smart environments. Definitions of smart environments may be taken into account as a first reference point, since smart products have to be considered in the context of their environment. For example, Das and Cook [6] define a smart

 ISBN: 978-1-61208-228-8

environment as the one that is able to acquire and apply knowledge about an environment and adapt to its inhabitants in order to improve their experience in that environment. It is noticed that the knowledge aspect has been recognized as a key issue in this definition. Mühlhäuser [2] refers to smart product characteristics that are attributed to future smart environments, i.e., "integrated interwoven sensors and computational systems seamlessly embedded in everyday systems and tools of our lives, connected through a continuous network." In this respect, smarter products can be viewed as those products that facilitate daily tasks and augment everyday objects. In 2007, AMI identified two motivating goals for building smart products [7]:

1) Increased need for simplicity in using everyday products, as their functionalities become ever more complex. Simplicity is desirable during the entire life-cycle of the product to support manufacturing, repair, or use.

2) Increased number, sophistication, and diversity of product components (for example, in the aerospace industry), as well as the tendency of the suppliers and manufacturers to become increasingly independent of each other which requires a considerable level of openness on the product side.

Mühlhäuser [2] observed that these product characteristics can now be developed due to recent advances in information technology as well as ubiquitous computing that provides a "real world awareness" in these systems through the use of sensors, smart labels, and wearable, embedded computers. According to Mühlhäuser [2], product simplicity can be achieved with improved product to user interaction (p2u). Furthermore, openness of a product requires an optimal product to product interaction (p2p). Knowledge intensive techniques enable better p2p interaction through self-organization within a product or a group of products. Indeed, recent research on semantic web service description, discovery, and composition may enable self-organization within a group of products and, therefore, reduce the need for top-down constructed smart environments [8]. Smart products also require some level of internal organization by making use of planning and diagnosis algorithms as stated by [2]:

> "A Smart Product is an entity (tangible object, software, or service) designed and made for self-organized embedding into different (smart) environments in the course of its lifecycle, providing improved simplicity and openness through improved p2u and p2p interaction by means of context-awareness, semantic self- description, proactive behavior, multimodal natural interfaces, AI planning, and machine learning."

Major characteristics of smart products are illustrated by comparing their essential features. For example, [9] define six major characteristics for smart products illustrated in Table 1 below. Table 2 provides a comparative presentation of the main characteristics of smart products. These characteristics include the following:

- Context-awareness - the ability to sense context
- Proactivity - the ability to make use of this context and other information in order to proactively approach users and peers
- Self-organization - the ability to form and join networks with other products.

In addition to the above characteristics, Mühlhäuser [2] and SPC emphasize the fact that smart products should support their entire life-cycle. In addition, special care should be devoted to offering multimodal interaction with the potential users, in order to increase the simplicity characteristics of the products.

TABLE 1. SMART PRODUCTS CHARACTERISTICS [9]

Characteristic	Description
Personalization	Customization of products according to buyer's and consumer's needs.
Business-awareness	Consideration of business and legal constraints.
Situatedness	Recognition of situational and community contexts.
Adaptiveness	Change product behavior according to buyer's and consumer's responses to tasks.
Network ability	Ability to communicate and bundle with other products.
Pro-activity	Anticipation of user's plans and intentions.

TABLE 2. A COMPARISON OF SMART PRODUCT'S CHARACTERISTICS [7]

Maass and Varshney [9]	Mühlhäuser [2]	Smart Products consortium [4]
Situatedness	Context-aware	Situation- and context-aware
Pro-activity	Proactive Behavior	Proactively approach the user
Network ability	Self-organized embedding	Self-organized embedding in smart product environments
	Support the entire life-cycle	Support the user throughout whole life-cycle

Maass and Varshney [9]	Mühlhäuser [2]	Smart Products consortium [4]
Personalization Business-awareness Adaptiveness	Multimodal Natural Interfaces	Multimodal interaction Autonomy Support procedural knowledge Emerging knowledge Distributed storage of knowledge

II. SOCIAL NETWORKING FOR SMARTER PRODUCTS AND SERVICES DESIGN

Communication of ideas, as a core for effective education and collaborative design, is the basis of distance and virtual learning. Social networks blend self-directed learning and prescribed, existing information. The self-directed element creates interest within a learner and the ability to access existing information facilitates its transfer, and eventual retention of knowledge acquired. There may also be a competitive element for educators to explore, since design activities are transparent in social networking. Ziegler [10] observed that social networking sites may radically change the educational system, since they offer the "capacity to motivate students as engaged learners", rather than what he considers the usual "passive observers of the educational process." However, there are also conflicting views in the literature regarding the usefulness of social networks in education and design. In today's interconnected world, social networking provides a great source of information and knowledge sharing that has not yet been fully explored to support collaborative products design and education.

Selwyn [11] performed an observational study of a group of students' online interactions with Facebook™ in the UK. Though the author cited many limitations of the study, some interesting findings included an observation that the social network site did not serve a meaningful role in making new partnerships. Rather, it maintained strong links already established in an emotionally close-knit group of people. Social networks share many functional elements with blog, a term coined recently as a shortened form of "web log," describing a page that is frequently updated with comments, links, images, and other media pertaining to a given subject. The blog makes a statement and offers a space below for readers to comment and respond. Social networks have taken the blog concept and applied it to a directory concept. People who are "linked" together can receive updates from others micro-blog inputs. The concept of social networking can be extended to collaborative design and modeling as means of facilitating team work and sharing product design experience in order to enhance team learning process, including collaborative online discussions, idea generation, peer review activities, and even debate [12,13].

III. SOCIAL NETWORKING IN EDUCATION

The proliferation of broadband-enabled interactive devices, such as smart cellular phones and media players, with social networking gadgets and application allows social communication and collaborative education activities to occur outside of lecture times. Another offshoot of social networking and blogging sites is that of wiki articles and their massive compilation, Wikipedia. A traditional understanding of an academic resource that "anyone can edit" seems unreasonable. The seemingly micro-managed and endlessly peer reviewed "live" nature of the document made Wikipedia a compelling new way to create, store, and integrate vast stores of knowledge [14]. The pull of social networking technologies cannot be ignored, as they have attracted millions of users in a short amount of time since their introduction. Ahram et. al. [23, 24, 25, 26] cited that the shortcomings such as the necessity of pre-existing offline relationships and, as of yet, unexploited educational and design opportunities may be addressed by serious initiatives and the integration of such technologies into modern educational and design methods and practices. The accessibility of these networks is more pervasive now than ever, thanks to gaming consoles and mobile devices. Prensky [15] claims that today's students "think and process information differently" from their pre-digital world counterparts. People born after the mid 1980s are part of a group of "digital natives" who take information technology and its use for granted. Today, data can be created anywhere and on a great variety of computing platforms. The ability to create and view data anywhere can translate to new learning opportunities. Several universities have already turned towards the web and outlets previously used only to sell music and video as a way to disseminate lecture materials. Apple Computer's iTunes TM software dominates the digital media player market. It has recently launched "iTunes University," a subset of its online media store devoted to distributing lectures and presentations from various academic institutions. All of these novel technologies and media distribution platforms offer unimagined learning opportunities. As of yet, the educational elements are largely unused compared to their strictly entertainment-related digital media. Many opportunities exist for providing students with this media but dissemination is not enough alone. Serious educational games, educator's involvement, and classroom activities sent to these services offering

 ISBN: 978-1-61208-228-8

interaction rather than "passive observation" would be valuable aids to the learning process. There is no doubt that today's traditional students consume more media and games than previous generations. They need only be given some structure and appropriate interactive learning media to augment their already media-enriched lives.

IV. Social Networking Systems Engineering Approach to Study Complex Human Behavior

The contemporary systems engineering process is an iterative, hierarchical, top down decomposition of system requirements [16]. The hierarchical decomposition includes Functional Analysis, Allocation, and Synthesis. The iterative process begins with a system-level decomposition and then proceeds through the functional subsystem level, all the way to the assembly and program level. The activities of functional analysis, requirements allocation, and synthesis will be completed before proceeding to the next lower level. SysML is a general-purpose visual modeling language for specifying, analyzing, designing, and verifying complex systems which may include hardware, software, information, personnel, procedures, and facilities (OMG SysML: http://www.omgsysml.org). SysML provides visual semantic representations for modeling system requirements, behavior, structure, and parametrics, which is used to integrate with other engineering analysis models [17, 23, 26].

Ahram et al. [23, 26] indicated that traditional machine learning techniques have some limitations for modeling human behavior, mainly the lack of any reference to the inherent uncertainty that human decision-making has. This problem can be partially solved with the introduction of Soft Computing (SC) to model human behavior via social networking applications. SC is an innovative approach to building computationally intelligent systems that differs from conventional (hard) computing in that it is tolerant of imprecision, uncertainty and partial truth. The guiding principle of soft computing is to exploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost. SC consists of several computing approaches, including neural networks, fuzzy set theory, approximate reasoning, and search methods, such as genetic and evolutionary algorithms. SC technologies provide an approximate solutions to an ill-defined problems encountered in social networking application and can help creating human behavioral models in an environment, such as during conflicts , in which users are not willing to give feedback on their actions and/or not able to fully define all possible interactions due to social and cultural barriers. Different techniques provide different capabilities to support the development of smarter products and services. For example, Fuzzy Logic provides a mechanism to mimic human decision-making that can be used to infer goals and plans; Neural Networks a flexible mechanism for the representation of common characteristics of a user and the definition of complex stereotypes; Fuzzy Clustering a mechanism in which a user can be part of more than one stereotype at the same time; and Neuro-Fuzzy systems a mechanism to capture and tune expert knowledge which can be used to obtain assumptions about the user.

Systems engineering teams along with product and service designers are responsible for verifying that the developed products and services meet all requirements defined in the system specification documents. The following procedures outline the relevant systems engineering process steps [18, 21, 26]:

- Requirements analysis: review and analyze the impact of operational characteristics, environmental factors, functional requirements and develops measures suitable for ranking alternative designs in a consistent, objective manner. Each requirement should be re-examined for consistency, desirability, applicability, and potential for improved return on investment [19]. This analysis verifies that the requirements are appropriate or develops new requirements for the smart product operation.

- Functional analysis - systems engineers and product designers use the input of performance requirements to identify and analyze system functions in order to create alternatives to meet system requirements. Systems engineering then establishes performance requirements for each function and sub-function identified.

- Performance and functionality - systems engineering allocates design requirements and performance to each system function. These requirements are stated in appropriate detail to permit allocation to software, systems components, or personnel. Performance and functionality allocation process identifies any special personnel skills or design requirements.

- Design Synthesis - designers and other appropriate engineering specialties develop a system architecture design to specify the performance and design requirements which are allocated in the detailed design. The design of the system architecture is performed simultaneously with the allocation of requirements and analysis of system functions. The design is supported with block and flow diagrams. Such diagrams support:

  - Identifying the internal and external interfaces
  - Permitting traceability to source requirements
  - Portraying the allocation of items that make up the design

 ISBN: 978-1-61208-228-8

  - Identifying system elements along with techniques for its test and operation
  - Providing a means for comprehensive change control management

- Documentation - the primary source for developing, updating, and completing the system and subsystem specifications. Smart product requirements and drawings should be established and maintained.

- Specifications - to transfer information from the smart product systems requirements analysis, system architecture design, and system design tasks. The specifications should assure that the requirements are testable and are stated at the appropriate specification level.

- Specialty engineering functions - participate in the systems engineering process in all phases. They are responsible for system maintainability, testability, producibility, human factors, safety, design-to-cost, and performance analysis to assure the design requirements are met.

- Requirements verification - systems engineering and test engineering verify the completed system design to assure that all the requirements contained in the requirements specifications have been met.

Model-based interactive human system approaches for design and modeling of smart systems and products differentiate between human performance and effectiveness criteria. These criteria determine a total system mission performance level and acceptability that is directly attributable to specific actions allocated to human performance metrics. These are indicators measure which performance effectiveness criteria are met [20, 21].

Currently there are few applications to facilitate human behavior modeling in social networking applications. One of the applications that support a full Human Systems Integration (HSI) within a systems engineering process is DOORS™ by Rational. DOORS or Dynamic Object Oriented Requirements System specifically tracks requirements for product or software design. Since the requirements process has many shared elements to knowledge management, DOORS facilitates requirements entry, organization into hierarchies, and display. Users make changes and link any requirement to sub requirements and related requirements. DOORS require individual users to have accounts. Each account can be restricted to elements of the database and given read-only or administrative-level rights. Changes made are tracked by user, allowing managers to trace changes down to the individual user level. IBM Rational DOORS [27] provides a structured framework for adding, viewing, and changing requirements.

V. HYBRID COMPUTATIONAL SOCIAL NETWORK ANALYSIS

The interest in Computational Social Network Analysis (CSNA) has been growing massively in recent years. Psychologists, anthropologists, sociologists, economists, and statisticians have given significant contributions, making it actually an interdisciplinary research area. This research summarizes the development of CSNA framework composed of methods used to (1) rapidly collect and (2) visualize socio-cultural network data in order to analyze economic and social data relationships between people, groups, organizations- and other knowledge-processing entities by integrating knowledge from available databases [23,26]. In particular, the first category (methods used for network data collection) aims to provide a dataset that helps study the effects that social networks have on different aspects of social and cognitive activities. To achieve this aim, the following methods were considered: (a) Socio-centric modeling: to examine sets of relationships between actors that are regarded for analytical purposes as bounded social collectives. (b) Ego-centric: to select focal actors (egos), and identify the nodes they are connected to.

The second category (methods used for network data visualization) aims to render data in easily understood graphical formats, thus making complex information usable and understandable by a broad community within the civilian economics and decision leaders to enable better decision-making at various levels (policy, and private sector sales operations) and to support cultural situational awareness for tactical decision-making and training. To achieve this aim, the following representations are used:

(a) Graphs: to visualize relationships among members of a narrow socio-cultural Network;

(b) Matrices: to visualize dense socio-cultural networks

(c) Maps: to manage a wide amount of data and information; and

(d) A mixed hybrid approach: to integrate different visualization perspectives according to the rapid socio-cultural assessment goals.

Different techniques provide different capabilities to support extensive cultural situational awareness analysis for tactical decision-making and rapid socio-cultural assessment. For example, techniques such as Self Organizing Maps (SOM) and Cellular Automata (CA) can be used to construct a user model by themselves or in combination with traditional machine learning techniques. Generated models based on available datasets can be classified in this work based on how human behaviors are represented as models,

 ISBN: 978-1-61208-228-8

and their purpose (Figure 1). To this end, two main dimensions are considered:

1) Granularity of the model: A model can be created for each individual (content-based modeling) or for clusters of users or a group (collaborative modeling)

2) Type of task for which the model is going to be used: Ahram et al. [26] define four basic types of tasks: (i) Filtering, (ii) Classification, (iii) Prediction, and (iv) Recommendation

The developed model supports prediction, which is the capability of anticipating future actions using past behavior [23,26]. A basic assumption is made with this approach: a user's immediate future is very similar to his/her immediate or midterm past, an approach used to describe and model the social network system emergent behavior as shown in Figure 1 [23, 26]; this is traditionally presented in the literature as content-based filtering [28]. Recommendation is the capability of suggesting interesting socio-cultural dataset elements relationships; Clustering is the process of locating "interesting" data and groups of interests from among the data. It is a technique that groups data with similar characteristics. The purpose of visualization using a Geographic Information System (GIS) is to map data presented in the models with SOM and CA onto a graphical representation to provide a qualitative idea of its properties. [26].

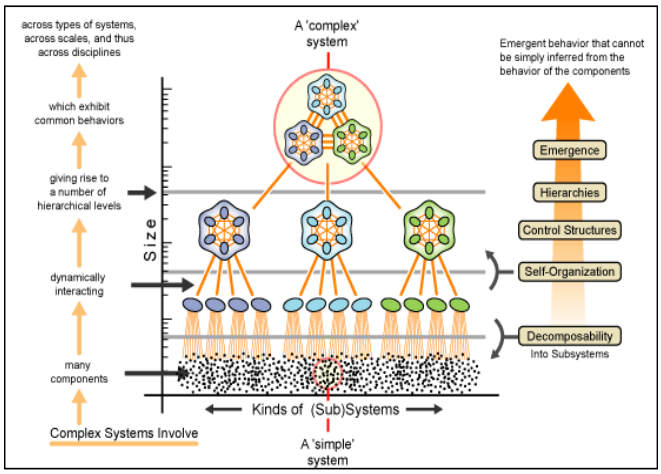

Figure. 1. Characterisitcs of a socio-economic cultural system to reflect emergent behavior

Self-Organizing Map's (SOM) is a neural network model for clustering and visualizing high-dimensional data. The SOM can be used to map high-dimensional complex socio-cultural data onto a low-dimensional space that is usually two dimensional. In general, data visualization is used as a way to aggregate large quantities of data (see Figure 2), and present them in a way that allows to: (1) quickly communicate rich messages (communication). (2) discover new, previously unknown facts and relationships (discovery). (3) getting better insight into things we already know (insight).

The visualization can be performed by using (i) graphs made up of nodes and connection lines and the numbers in each cell stand for specific relationships among these values, (ii) matrices where row and columns stand for actors and properties, (iii) maps, and/or (iv) a hybrid approach.

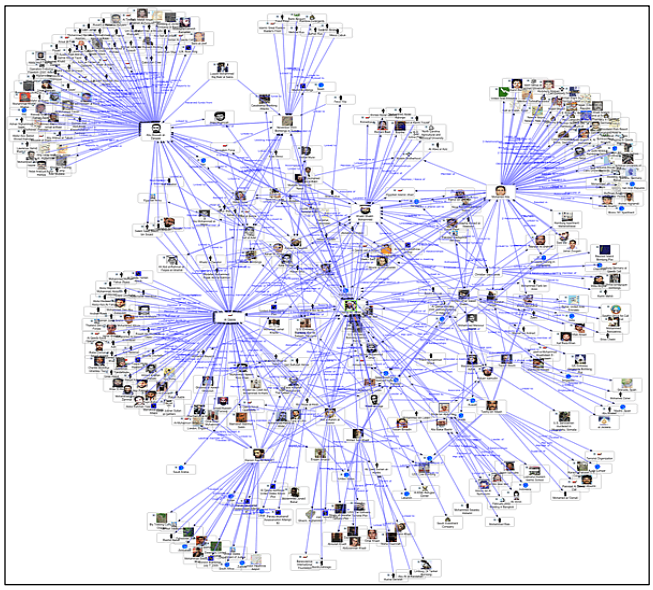

Figure.2. Example Visual Socio-Cultural Network Analysis

VI. CONCLUSION AND FUTURE WORK

As an introductory contribution to the application of social networking and systems engineering process for the design and development of smarter products and services, this paper provides a motivation and quest for integrated social networking approach to systems engineering and to study complex human behavior. While a large number of disciplines and research fields must be integrated towards development and widespread use of smarter products, considerable advancements achieved in these fields in recent years indicate that the adaptation of these results can lead to highly sophisticated yet widely useable collaborative social networking applications for smart products.

Future research will emphasize the applications of systems engineering and social-networking to design and modeling of specific services and will demonstrate the benefits in supporting and facilitating the understanding of complex human behavior and to better identify crucial user needs.

 ISBN: 978-1-61208-228-8

ACKNOWLEDGMENT

This research was sponsored by the Office of Naval Research (ONR) and the Office of Secretary of Defense (OSD) under contract No. N00014-11-1-0934. The authors acknowledge the helpful guidance by ONR Program Management, and the contributions of the technical team.

REFERENCES

[1] G. Allmendinger and R. Lombreglia, "Four Strategies for the Age of Smart Services", Harvard Business Review, 2005, 83(10):131-145.

[2] M. Mühlhäuser, "Smart Products: An Introduction. In Constructing Ambient Intelligence", AMI 2007 Workshop, pages 154 – 164, 2008.

[3] M. Weiser, "The computer of the 21st century", Scientific American, 1991, 265(3):66-75.

[4] J. Ahola, "Ambient Intelligence". ERICM News, 2001, (47).

[5] E. Arts and B. de Ruyter, "New research perspectives on ambient intelligence". Journal of Ambient Intelligence and Smart Environments, 2009, 1:5-14.

[6] S. Das and D. Cook, "Designing Smart Environments: A Paradigm Based on Learning and Prediction", In: Shorey, R., Ananda, A., Chan, M.C., Ooi, W.T. (eds.) Mobile, Wireless, and Sensor Networks: Technology, Applications, and Future Directions, 2006, pp. 337–358. Wiley, Chichester.

[7] M. Sabou, J. Kantorovitch, A. Nikolov, A.Tokmakoff, X. Zhou and E. Motta. "Position Paper on Realizing Smart Products: Challenges for Semantic Web Technologies", Report by Knowledge Media Institute, 2009, http://people.kmi.open.ac.uk/marta/papers/ssn2009.pdf form [retrieved: August, 2012].

[8] S. Chandrasekharan, "The Semantic Web: Knowledge representation and affordance", In: B. Gorayska, and J.L. Mey (eds). Cognition and Technology: Co-existence, Convergence, and Co-evolution, Amsterdam/Philadelphia: Benjamins, 2004, pp.153-174

[9] W. Maass and U. Varshney, "Preface to the Focus Theme Section: 'Smart Products'. Electronic Markets", 2008, 18(3):211-215.

[10] S. Ziegler, "The (mis)education of Generation M' Learning, Media and Technology", 2008, 32, 1, pp.69-81

[11] N. Selwyn, "Screw blackboard… do it on Facebook!: An investigation of students' educational use of Facebook", Paper presented to the Pole 1.0 – Facebook social research symposium, November 15, 2007 at University of London: http://www.scribd.com/doc/513958/Facebook-seminar-paper-Selwyn [retrieved: October, 2009].

[12] J. Davies, and G. Merchant, "Looking from the inside out: academic blogging as new literacy" in C. Lankshear, and M. Knobel, (eds.) 'A New Literacies Sampler' New York, Peter Lang, 2007.

[13] F. Dubet, "Dimensions and representations on student experience in mass universit", Revue Francaise de Sociologie, 2004, 35, 4, pp.511-532

[14] C. Madge, J. Meek, J. Wellens and T. Hooley, "Facebook, social integration and informal learning at university: 'It is more for socialising and talking to friends about work than for actually doing work", Learning, Media and Technology, 2004, Routledge.

[15] M. Prensky, "Digital Game-Based Learning", McGraw-Hill, New York, 2001, ISBN 0-07-136344-0

[16] D. K. Hitchins, "Systems Engineering: A 21st Century Systems Methodology. Chichester, UK: John Wiley & Sons, 2007.

[17] S. Friedenthal, A. Moore and R. Steiner, "A Practical Guide to SysML: The Systems Modeling Language", Morgan Kaufmann; Elsevier Science, 2008.

[18] Defense Acquisition University (DAU) Guidebook 2004, Chapter 4: Systems Engineering.

[19] T. Z. Ahram and W. Karwowski, "Measuring Human Systems Integration Return on Investment". The International Council on Systems Engineering – INCOSE Spring 2009 Conference: Virginia Modeling, Analysis and Simulation Center (VMASC), 2009, Suffolk, VA. USA.

[20] T. Z. Ahram, W. Karwowski, B. Amaba and P. Obeid, "Human Systems Integration: Development Based on SysML and the Rational Systems Platform", Proceedings of the Industrial Engineering Research Conference, 2009, Miami, FL. USA.

[21] Systems Engineering Fundamentals, Defense Acquisition University Press, 2001.

[22] W. Karwowski and T. Z. Ahram, "Interactive Management of Human Factors Knowledge For Human Systems Integration Using Systems Modeling Language". Special Issue for Information Systems Management, Journal of Information Systems ManagemenT, 2009, Taylor and Francis.

[23] T. Z. Ahram, W. Karwowski, and B. Amaba, "Collaborative systems engineering and social networking approach to design and modeling of smarter products". Behav. Inf. Technol. 30, January, 2011, 13-26.

[24] T. Z. Ahram and W. Karwowski, "Social Networking Applications: Smarter Product Design for Complex Human Behavior Modeling", Human Computer Interaction Proceedings. 14th International Conference on Human-Computer Interaction (HCII 2011), Orlando, FL, July 9-14, 2011.

[25] T. Z. Ahram, W. Karwowski, " Visual Social Network Analysis: Effective Approach to Model Complex Human Social, Behavior & Culture". The 18th World Congress on Ergonomics (IEA) held February 12-16, 2012, in Recife, Brazil.

[26] T. Z. Ahram and W. Karwowski, "Developing Human Social, Cultural, Behavior (HSCB) Ontologies: Visualizing & Modeling Complex Human Interactions", Presented at the Office of Secretary of Defense Human Social, Culture Behavior Modeling (HSCB Focus 2011), Chantilly, VA, February 8-10, 2011

[27] IBM Rational DOORS software website: http://www.ibm.com/software/awdtools/doors/ [retrieved: September, 2012].

[28] M. Balabanovic and Y. Shoham, "Learning Information Retrieval Agents: Experiments with Automated Web Browsing". In Working Notes of the AAAI Spring Symposium Series, 1995, pp. 13-18. Stanford University.

 ISBN: 978-1-61208-228-8

An Efficient Computation of Reachability Labeling for Graph Pattern Matching

Bhargavi Balla, Supreethi Kalyandurgam Pujari
Department of Computer Science & Engineering
JNTUH College of Engineering
Hyderabad, India
Emails: {bhargavi.bbv1, supreethi.pujari}@gmail.com

***Abstract*— Due to the rapid growth of Internet, most of the data that is available in the Internet that is archived/ analyzed, is graph structured in nature as graphs form a powerful modeling tool. The problem of graph pattern matching is to find all the tuples that match a user-given graph pattern from a large directed graph. For faster access of paths in the large directed graph, transitive closure of the graph is compressed and maintained using 2-hop reachability labeling technique by assigning every node a 2-hop label. These 2-hop labels are computed using a geometry-based approach that will be useful in solving the graph pattern matching problem. In this paper, a geometry-based approach that computes the 2-hop reachability labeling is described. The experimental results show that the proposed approach efficiently computes the compressed transitive closure technique of reachability labeling.**

***Keywords*- graph pattern; graph matching; 2-hop cluster; 2-hop labeling; 2-hop cover.**

I. INTRODUCTION

Graphs form a powerful modeling tool to represent various networks in different areas like chemistry, biology, web, etc. In online social networking systems like Facebook and Twitter, the relationships among users and their proximity can be conveniently expressed using graphs. Thus, there is a demand for efficiently querying the graph data.

Graph database [1] is a large labeled directed graph or a collection of labeled directed graphs. A graph pattern is a sequence of nodes and edges which is constructed by connecting nodes based on links/relationships between them as required by the user. Given a graph database and a graph pattern, finding all the set of tuples (an ordered sequence of vertices) that match a user given graph pattern is the graph pattern matching problem. For instance, in analyzing online social networking systems, a large graph can be obtained where the job-title attribute on each node can be regarded as label. A small graph pattern can be to discover connections between several people with specified jobs. But, the graph pattern matching problem is challenging as graph data can be large and graph patterns can be large and complex.

To access the paths in a large graph data faster, its compressed transitive closure is pre-computed using 2-hop reachability technique which involves assigning a graph code termed 2-hop reachability label to each node of the directed graph. The computation of 2-hop reachability labeling for the graph is found to be NP-hard [3]. In this paper, a geometry-based approach is implemented to efficiently compute the 2-hop reachability labels for a large directed graph which is a nearly optimal solution. The graph codes computed will be useful in solving the graph pattern matching problem in relational database context [1].

Section II covers the related work done for finding the efficient techniques to solve the problem of graph pattern matching and 2-hop reachability labeling. Section III describes the prominent compressed transitive closure techniques while section IV describes the procedure to compute 2-hop reachability labels efficiently. Section V reveals the experimental results and analysis and in section VI, we conclude the paper with future work.

II. RELATED WORK

Extensive survey has been done for finding efficient techniques to solve the graph pattern matching problem [10]. It includes the survey on tree-pattern matching techniques [8][9], graph pattern matching techniques [1][5] and extensive survey on multi-interval encoding [4] and 2-hop labeling [2][3][6][7].The problem of tree pattern matching is to find the set of patterns from a large tree that match the given tree pattern. Bruno et al. [8] used stack encoding scheme for tree pattern matching in XML documents with elements and parent-child relationships rendering it as tree. Chen et al. [9] further improved by using hierarchical stack encoding scheme for tree pattern matching. But, these techniques do not work on graph data directly as graphs do not have the good acyclic property of trees.

For faster access of paths and for testing if two nodes are reachable, transitive closure is pre-computed and stored in compressed form. Multi-interval encoding defined by Agrawal et al. [4] is a compressed transitive closure technique used for faster processing of graph pattern matching of graph-structured documents in [5]. The 2-hop reachability labeling defined by Cohen et al. [3] is a compressed transitive closure technique where each node is assigned labels that represent the reachability information of the node. The problem of 2-hop cover is to find the minimum size of 2-hop cover for a given graph, which is proved to be NP-hard [3]. Cohen et al. [3] show that a greedy algorithm exists to compute a nearly optimal solution for the 2-hop cover problem. The resulting size of

 ISBN: 978-1-61208-228-8

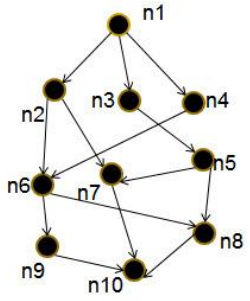

Figure 1. A Directed Graph G

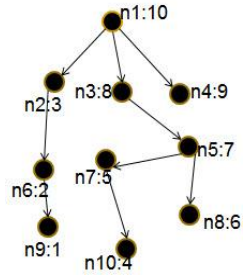

Figure 2. Optimum tree cover with postid for each node

the greedy algorithm is larger than the optimal by atmost O(log n). The basic idea is to solve the minimum 2-hop cover problem as a minimum set cover problem [3]. Schenkel et al. [6][7] implemented divide and conquer approach where the directed graph is partitioned into subgraphs and 2-hop cover is computed for each partition as in [3] and then the results are combined. The approach developed by Cohen et al. focused on finding minimum overlap among the subsets for finding 2-hops of the graph while J. Cheng et al. [2] implemented a geometry-based approach and focused on finding the minimum number of subsets which finds the 2-hops faster than the former approach.

The survey [10] resulted in discovering an efficient geometry-based approach [2] for finding reachability labels of the directed graph.

III. COMPRESSED TRANSITIVE CLOSURE COMPUTATION TECHNIQUES

Transitive closure represents all the set of paths between the nodes of the graph that satisfy transitivity property. The transitive closure size is defined as the total number of paths present in the transitive closure. By pre-computing transitive closure, we can access shortest paths faster and check the existence of paths between the two nodes. But, the transitive closure size is very large compared to the total number of vertices/edges that represent the directed graph. The following are the two different techniques to compute the compressed transitive closure efficiently.

A. *Multi-interval Encoding*

Multi-interval encoding technique involves assigning every node a postid and interval list with atleast one interval that together represent the compressed transitive closure of the directed graph.

Initially, the directed graph is converted into DAG (directed acyclic graph) by computing the maximal strongly connected components of the directed graph and assigning a node to represent each maximal strongly connected component in DAG. Then, multi-intervals are computed using Agrawal et al. algorithm [4]. To compute multi-intervals, an optimum tree cover is derived first from the DAG [4]. Then, postids (the numbers) are assigned to the nodes in post-order traversal order of the optimum tree.

For instance, consider the optimum tree cover shown in Fig. 2 of the directed graph G in Fig. 1. In Fig. 2, consider a node "n9" to which the assigned postid is 1, which is assigned based on the post-order traversal of the optimum tree. To each node of the optimum tree cover, an interval [s, e] is assigned where 'e' is the postid of the current node and's' is the postid of the lowest descendant node.

For each leaf node 'v' with postid 'i', its interval assigned is [i, i] and that of its parent node with postid 'j' is [i, j]. For instance, from Fig. 2, the interval of the leaf node "n9" is [1, 1] and interval of its parent node "n6" is [1, 2] (shown in Table I). Multiple intervals for each node of the graph come into existence if there are back edges to the nodes in the DAG. For instance, in Fig. 1, there is a back edge (n8, n10) for node "n8" hence the interval of node "n10", i.e. [4, 4] is added to the interval list of "n8" as shown in Table I.

In general, there exists a path a~b if and only if the postid of 'b' lies in atleast one interval in the interval list of 'a'. For instance, there exists a path n3~n8 as the postid of "n8" is 6 which lies in the interval [4, 8] of n3. Multi-intervals for the other nodes of the directed graph are same

TABLE I. MULTI-INTERVAL ENCODING OF G

V	pid	I
n1	10	[1, 10]
n2	3	[1, 6]
n3	8	[4, 8]
n4	9	[[1, 2], [4, 4], [6, 6], [9, 9]]
n5	7	[4, 7]
n6	2	[[1, 2], [4, 4], [6, 6]]
n7	5	[4, 5]
n8	6	[[4, 4], [6, 6]]
n9	1	[[1, 1], [4, 4]]
n10	4	[4, 4]

 ISBN: 978-1-61208-228-8

as the multi-intervals of the maximal strongly connected nodes of the directed graph that represent them in DAG. Table I shows the multi-interval encoding of each node of G shown in Fig. 1. Thus, multi-interval encoding represents the compressed transitive closure of the directed graph.

The disadvantage of multi-interval encoding is that, it is lengthy and its storage cost increases with increase in the number of vertices/edges of the directed graph and sorting is required to perform graph pattern matching using the multi-interval encoding technique. Hence, the more efficient 2-hop reachability labeling or 2-hop cover technique is opted to compute and store the compressed transitive closure.

B. 2-hop Reachability Labelling

A hop in a directed graph is defined by a path in the graph and one of the end points of the path. For each node 'v' in a directed graph G (V, E), a label $L(v) = \{L_{in}(v), L_{out}(v)\}$ is assigned where $L_{in}(v)$ represents the set of the nodes in G that can reach 'v' & $L_{out}(v)$ represents the set of nodes in G that are reachable from 'v', (hence the name 2-hop) which define the 2-hop reachability labeling [3].

A 2–hop cover is a 2–hop labeling of directed graph G such that if there is a path u~v in G, then $L_{out}(u) \cap L_{in}(v) \neq \emptyset$. 2-hop cover is computed such that the transitive closure of graph is covered.

Table II shows the 2-hop reachability labels for the nodes of the directed graph G of Fig. 1. For instance, L (n3) = {{n1}, {n5}} is the 2-hop reachability label for a node 'n3' where $L_{in}(n3) = \{n1\}$ & $L_{out}(n3) = \{n5\}$. There exists a path n3~n8 as $L_{out}(n3) \cap L_{in}(n8)$ is {n5}. Thus, 2-hop reachability labelling represents the compressed transitive closure of the directed graph.

The problem of finding 2-hop cover is to assign the 2-hop reachability labels such that the total size is minimum which is found to be an NP-hard problem as it can be reducible to minimum set-cover problem which has no optimal solution. Minimum set cover problem is to find the subsets with minimum overlap covering all the paths. Each subset has a center w associated with it represented as $S(F_w, w, T_w)$ which is termed the 2-hop cluster with center 'w' and F_w and T_w constitute all the set of nodes that are reachable from 'w' and that can reach 'w' respectively. The center with maximum cost is selected. The cost is assigned to the center node based on the criterion of maximum number of paths that the node can cover.

TABLE II. 2-HOP REACHABILITY LABELING OF G

V	Lin(v)	Lout(v)
n1	{}	{n5, n6}
n2	{n1}	{n6}
n3	{n1}	{n5}
n4	{n1}	{n6}
n5	{}	{n1, n3}
n6	{}	{}
n7	{n2, n5}	{n10}
n8	{n5, n6}	{n10}
n9	{n6}	{n10}
n10	{n5, n6}	{}

IV. EFFICIENT COMPUTATION OF 2-HOP LABELLING FROM MULTI-INTERVAL ENCODING USING GEOMETRY-BASED APPROACH

We computed 2-hop clusters by implementing the geometry-based concept behind the algorithm in [2]. The geometry-based technique involves the following steps.

1. Construction of virtual reachability map.
2. Computation of rectangular map for each node 'w' that forms a bipartite graph with center 'w'.
3. Derivation of 2-hop cluster by mapping the rectangular map that has maximum matches with the virtual reachability map.

The steps 2 and 3 are repeated until the virtual reachability map is completely covered. The set of 2-hop clusters together represent the 2-hop cover.

Initially, the directed graph is converted to DAG and the reverse DAG is constructed. Reverse DAG is constructed by reversing the direction of edges of the DAG. Let "I" be the set having multi-interval encoding information of the nodes of DAG, i.e., for each node, postid and interval list is stored in "I". Let "It" be the set having multi-interval encoding information for the nodes of reverse DAG. Let $p(v_i)$ and $pt(v_i)$ be the postids of the node v_i in I and It respectively & let $I(v_i)$ & $It(v_i)$ denote the interval list of node v_i in "I" & "It" respectively where v_i is one of the vertices in V of the DAG G(V, E). The following pseudocode implements the steps of geometry-based approach and computes the 2-hop cover efficiently.

Algorithm 2-HopCover (I, It, V)

```
{ //size(V) returns the number of vertices of DAG in set V.
        n:=size(V);
        for i:=1 to n
        {
                for j:=1 to n, v_i≠v_j
                {
                for k:=1 to size(I(v_i)) //[x_k, y_k] is in I(v_i)
                if (x_k<=p(v_j)<=y_k) f[p(v_j)][pt(v_i)]:=1;
                }
        }//virtual reachability in 2D array f
        do
        {
        for m:=1 to n
       {m2:=Rect(v_m); if(max<m2) {max:=m2; w:=v_m; }}
```

```
//get w in V with maximum count from Rect(w).Let rf[i][j]
//be the array where reachability of node 'w' is stored.
for i:=1 to n{ for j:=1 to n{if(rf[i][j]=f[i][j] & f[i][j]=1){
F.add (p^-1(i)); T.add (pt^-1(j)); f[i][j]:=0;}}}
H.add(S (F, w, T)); F.empty (); T.empty ();
} while (atleast one value in 2D array f is 1);
return H;
}
Rect(w)
{ count:=0;
//I(w)={[s1,e1],s2,e2]..[sn,en]} & similarly for It(w).
        for k:=1 to size(I(w)) for i:=s_k to e_k
        for l:=1 to size(It(w)) for j:=s_l to e_l
        {
                rf[i][j]:=1; rf[p(w)][pt(w)]:=0;
                if(rf[i][j]=f[i][j]=1) count++;
        }
return count;
}
```

Pseudocode to compute 2-hop cover

In the pseudocode, the multi-interval encoding information is taken as input and the output returned is the 2-hop cover H. The explanation of the pseudocode along with the steps of the geometry-based technique is given below.

A. Construction of Virtual Reachability Map

For every node v_i, its reachability information is stored in a 2D array 'f' in the pseudocode defined as follows:

f[i][j]=1 if postid 'i' of a node ($p[v_j]$) lies in one of the intervals of v_i, and 'j' is the postid of the current node v_i ($pt[v_i]$) in reverse DAG.

This 2D array is termed as the virtual reachability map. This virtual reachability map contains the complete reachability information of the DAG.

B. Computation of Rectangular Map

Then, a rectangular map is created for each node v_i and stored in a temporary 2D array "rf". This map is created from the interval lists I and It of the current node v_i. Let I= [[s1, e1], [s2, e2],.[sk, ek]..[sn, en]]. Let It = [[$s1^1$, $e1^1$], [$s2^2$, $e2^2$],. .[sl, el]... [sn^n, en^n]] for v_i. The rectangular map "rf" computed in Rect(w) in the pseudocode is defined as follows:

For each interval [sk, ek] in I,
For each interval [sl, el] in It,
rf[i][j]=1 for all integers i such that i is in [sk, ek] and for all integers j such that j is in [sl, el].
rf[i][j]=0 if $p[v_i]$=i and $pt[v_i]$=j.

C. Derivation of 2-hop Cover

The procedure Rect(w) returns the total number of matching 1s of the virtual reachability map with "rf" of the node 'w' which is stored in variable "count". The node which has maximum value of "count" is selected. From the matching 1's of "rf" of such node 'w', a 2-hop cluster S (F, w, T) is derived. For each matching value, i.e. for each f[i][j]=rf[i][j]=1, where 1<=i<=n and 1<=j<=n, add the node v_i to F which has $p[v_i]$=i and add the node v_j to T which has $pt[v_j]$=j. Each 2-hop cluster formed can be visualized as a bipartite graph with 'w' as the center node of the bipartite graph. Thus, a 2-hop cluster is created which is added to 2-hop cover H. Then, assign to all the matching 1s in the virtual reachability map the value 0 and remove 'w'. This process of matching the rectangular maps with the virtual reachability map is repeated until no value in the virtual reachability map is 1. Thus, the 2-hop cover of DAG is computed.

The 2-hop clusters computed for the directed graph in Fig. 1 implementing the above pseudo-code are {[n1, n2, n4], n6, [n10, n8, n9]}, {[n1, n3], n5, [n10, n7, n8]}, {[n7, n8, n9], n10, []}, [], n1, [n1, n2, n3, n4]} & {[], n2, [n7]}. From the 2-hop cover, 2-hop labels can be derived for the nodes of the DAG. For each cluster S(Fw, w, Tw) , for each node 'u' in Fw, add 'w' to Lout (u) and for each node 'v' in Tw, add w to Lin (v).For instance, 2-hop labels derived for the directed graph in Fig. 1 are shown in Table II. Thus, 2-hop labels are constructed for the nodes of DAG. 2-hop labels for the other nodes of the directed graph are same as 2-hop labels of the maximal strongly connected nodes of the directed graph that represent them in DAG.

V. EXPERIMENTAL RESULTS

The graph data can be a large real XML document like DBLP data (http://dblp.uni-trier.de/xml/), or synthetic XML data like XMark [11] which are parsed to derive the directed data graph. XMark is a synthetic XML benchmark known for its irregular schema. There are elements that internally refer to other elements in the document which can be used to encode the XMark XML document as a directed graph with elements as nodes and parent-child relationships and referencing relationships as edges. The xmlgen tool of XMark constitutes a scaling factor which can be adjusted for generating XML documents with varying sizes.

We conducted our experiments on Dell laptop with 2.10 GHz processor and 3.0GB RAM running on Windows 7 in Java. 2-hop clusters are computed for XMark benchmark XML files using the efficient geometry based approach with the results outlined in the Table III. The last column in the Table III shows 2-hop labels which are computed from the 2-hop clusters (shown in third column of Table III). Fig. 3

 ISBN: 978-1-61208-228-8

TABLE III. EXPERIMENTED DATA AND RESULTS

XMark scaling factor	|V| of Directed Graph	|E| Of Directed Graph	No. of 2-hop clusters |S|	2-hop label size |L|
0.0001	372	438	109	891
0.0005	757	882	182	2444
0.0007	1143	1355	304	3237
0.0008	1158	1358	292	3419
0.001	1677	1961	480	4663

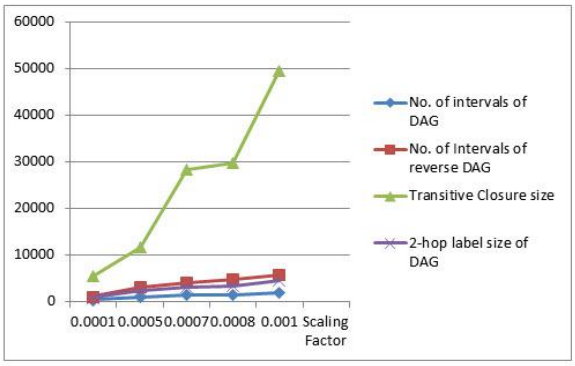

Figure 3. Comparison of Multi-interval Encoding size and 2-hop Labeling size to Transitive Closure size

shows significantly less 2-hop label size computed using our approach when compared to transitive closure size.

VI. CONCLUSION AND FUTURE WORK

The approach for computing 2-hop clusters is an efficient geometry-based approach which is tested on XMark XML files. From each 2-hop cluster $S(F_w, w, T_w)$, 2-hop reachability labels are constructed by adding center node 'w' to the label Lin(v) where 'v' is one of the nodes in T_w and adding center node 'w' to the label Lout(u) where 'u' is one of the nodes in F_w. The results shown in Fig. 3 of 2-hop labels computed show the significant amount of compression of transitive closure which indicates the efficiency of our approach. Using the 2-hop labels computed from 2-hop clusters, base relations and cluster-based index will be constructed that will be used in implementing the join-based algorithms [1] for efficiently solving the graph pattern matching problem.

REFERENCES

[1] J. Cheng, J. Xu Yu and P.S. Yu , "Graph Pattern Matching: A Join/Semijoin Approach", IEEE Transactions on Knowledge and Data Engineering, Vol. 23, No. 7, Jul. 2011, pp. 1006-1021, doi: 10.1109/TKDE.2010.169.

[2] J. Cheng, J. Xu Yu, X. Lin, H. Wang and P.S. Yu, "Fast Computation of Reachability Labeling for Large Graphs", Proc. International Conference of Extending Database Technology: Advances in Database Technology (EDBT '06), 2006.

[3] E. Cohen, E. Halperin, H. Kaplan and U. Zwick, "Reachability and Distance Queries via 2-Hop Labels", Proc. ACM-SIAM Symp. On Discrete Algorithms, (SODA 02), Jan. 2002, pp. 937-946.

[4] R. Agrawal, A. Borgida and H.V. Jagadish, "Efficient Management of Transitive Relationships in Large Data and Knowledge Bases", Proc. ACM Int'l Conf on Management Of Data, Jun. 1989, pp. 253-262.

[5] H. Wang, W. Wang, X. Lin and J. Li , "Coding-based Join Algorithms for Structural Queries on Graph-Structured XML Document", Vol. 11, No. 4, Springer 2008, pp. 485-510.

[6] R. Schenkel et al., "HOPI: An Efficient Connection Index for Complex XML Document Collections", Proc. Int'l Conf.(EDBT '04), 2004, pp. 237-255.

[7] R. Schenkel, A. Theobald and G. Weikum, "Efficient Creation and Incremental Maintenance of the HOPI Index for Complex XML Document Collections", Proc. 21st Int'l Conf. Data Eng. (ICDE '05), Apr. 2005, pp. 360-371.

[8] N. Bruno, N. Koudas and D. Srivastava, "Holistic Twig Joins: Optimal XML Pattern Matching", Proc. ACM SIGMOD 2002, pp. 310-321.

[9] S. Chen et al., "Twig^{2}stack: Bottom-Up Processing of Generalized-Tree-Pattern Queries over XML Documents", Proc. Int'l conf. on Very Large DataBases,(VLDB'06), 2006, pp. 283-294.

[10] B. Bhargavi and K.P. Supreethi, "Graph pattern mining: A survey of issues and approaches", International Journal of IT and Knowledge Management, Vol.5, No. 2, pp. 401-407, Jul. 2012.

[11] A. Schmidt, F. Waas, M. Kersten, M.J. Carey, I. Manolescu and R. Busse, "XMark: A Benchmark for XML Data Management," Proc.28th Int'l Conf. Very Large Data Bases (VLDB '02), 2002.

 ISBN: 978-1-61208-228-8

Everywhere Race!: A Social Mobile Platform for Sport Engagement and Motivation

Fabrizio Mulas *, Paolo Pilloni*, Salvatore Carta*
*Dept. of Mathematics and Computer Science, University of Cagliari via Ospedale 72 09124 Cagliari, Italy
Email: {fabrizio.mulas, paolo.pilloni, salvatore}@unica.it

***Abstract*—A lot of medical researches demonstrate that in the last decades there has been a serious increase of chronic diseases such as obesity, diabetes and correlated pathologies. These studies emphasize the key role of a healthy lifestyle in the prevention and cure of these kind of health problems. A healthy lifestyle can be broadly summarized in terms of regular physical activity, a balanced diet, a positive outlook and absence of harmful substances such as alcohol and tobacco. Researchers around the world are experimenting with new techniques and technologies to support and motivate people in their daily physical activity. Ubiquitous mobile devices and their intrinsic relationship with social networks are becoming one of the most popular persuasive technologies adopted in this research area. In this paper we present and evaluate an innovative mobile Android application, called *Everywhere Race!*, designed to motivate people in a wide range of speed-based sport activities. The peculiarity of the application consists in its groundbreaking way to engage people in sports. Thanks to an intensive usage of social networks and its fun-oriented design it makes it possible, for the first time ever, to compete in real-time with people from all over the world as in a real competition. A preliminary evaluation of the application, conducted on 35 users, shows that the application motivates both non-habitual sportsmen to start working out and habitual ones to keep on exercising.**

Keywords-Sport; Healthy Lifestyle; Motivation; Engagement; Gamification; Serious Game; Human Computer Interaction.

I. INTRODUCTION

Thousands of medical studies have put in evidence the benefits of regular physical activity. These studies demonstrate that a healthy lifestyle is a predominant factor for people well-being. It helps in the prevention of several chronic diseases and it is closely related to a wide array of mental health benefits (see for example, [2][3]).

Despite these scientific results, millions of people worldwide conduct a sedentary lifestyle. Data collected during the years 2008-2010 by the World Health Organization (WHO) reveal that we are living in a sedentary society. These statistics speak clearly [4]:

- In 2008 1.4 billion adults were overweight
- Nearly 40 million children are overweight
- 65% of world population lives in countries where overweight and obesity kill more people than underweight

WHO suggests some guidelines to be followed in order to overcome or prevent health problems. Here are some proposed tips:

- Engage in regular physical activity
- Limit energy intake from total fats
- Increase consumption of fruit and vegetables, as well as legumes, whole grains and nuts

In the last years, a conspicuous number of research teams and famous sport brands were and still are involved in studying and developing new techniques to motivate people to pursue a more active lifestyle. Many of these studies share a common approach: the exploit of new technologies, such as smartphones and social networks, as primary means of promoting an active lifestyle especially among young people.

As stated by Batussi et al. [5] the products developed by these researches can be grouped in three main categories:

- Computer-supported physical games
- Virtual trainers
- Mobile applications and devices

The last category seems to be the most promising one if we consider both the high number of related researches and the solutions developed. One key success factor is certainly due to the high potentialities of new technologies.

Mobile devices are highly portable and this allows users to be assisted everywhere and every time they need it. For example, Mulas et al. [17] exploit a mobile application that behaves as a virtual personal trainer for supporting and motivating people in their running activities.

Most of the proposed applications, however, try to motivate users focusing mainly on people performance rather than the social and ludic aspects of sport. In some cases this can be a limiting factor especially when the aim is to attract non-habitual sportsmen.

Everywhere Race! (EWR) has been designed to overcome these limitations. Its main purpose is to allow users from all over the world to compete against each other in different speed-based sports.

The application, in a completely different way than other existing applications, allows real-time virtual competitions among participants.

A virtual race is similar to a real one in the sense that it has a priori known start time and distance and both are the same for all the participants, no matter where they are.

The application is designed to promote a strong social interaction among users by means of the famous social network Facebook.

Indeed a user can easily check all races where its friends are enrolled in and join them. Alternatively, it can create a

 ISBN: 978-1-61208-228-8

new race and, perhaps, invite its friends. Furthermore, EWR highlights friends activities, such as their last race results and classifications, in order to foster social engagement.

Once a user chooses the race that best meets his preferences, he only has to wait till the scheduled start time. A few minutes before it, EWR begins a countdown for the user, so that all participants will start racing at the same time, wherever they are in the world.

During the competition the application records and shows to the user his race data – such as the covered distance so far, the average speed/pace and his current position in the competition – and those of their opponents in terms of position and distance from him as in a real event.

When all participants finish the race, the application establishes and publishes the final classification, with arriving positions, times and speeds, as in a real race (Figure 6).

If the user wants, his result will be posted on his Facebook wall.

Everywhere Race! offers another innovative feature: it is even possible to associate to a real sport event (such as a marathon) a virtual one. In this way people that can not be present at the real event can virtually participate and compete against all other participants. For example, imagine to take part at the New York Marathon without being physically there, perhaps running in your city park or in your treadmill (indeed modern smartphone are equipped with accelerometers that can be exploited to realize a pedometer).

Promising results coming from the relatively new serious game research area (see for example, [7][8]) suggested us to develop the application leveraging the importance of fun and social relationships as a means to promote the sport in as many people as possible.

With respect to the state-of-the-art solutions, our application offers users the opportunity to compete in real-time virtual races that are very similar to the real ones. At the same time the deep integration with Facebook community makes sport practice even more engaging. These factors help to push people toward a healthy spirit of competition that is a well known strong motivation element.

Other existing solutions do not offer users the possibility to perform real-time competitions with other people. The most common case is to allow a user to compete against his past performances as a means to improve himself. To the best of our knowledge there is only one application, named Softrace [23], similar to *Everywhere Race!*. The application allows users to take part in a virtual real-time competition. Anyway it differs from EWR mainly because it does not follow the common rules of a real race (see Section II).

The rest of this paper is organized as follows: Section II surveys the state of the art in the field of pervasive computing. Section III details the application design while section IV reports preliminary test results. Section V concludes the paper.

II. Related Work

In this section, with no claim to being exhaustive, we will report some examples of technological systems developed to support people during physical activities.

Hoysniemi [9] illustrates some results regarding the use of the famous dance video game Dance Dance Revolution. In this paper the author investigates the positive effects of gaming with regard to motivational, physical and social factors.

Jayant et al. [10] studied the effects of using human movements as game controller. To achieve this they have developed MarioFit, a system to play the Nintendo game Mario Bros on a PDA.

IJsselsteijn et al. [11] propose a study on intrinsic motivation enhancement. The research is based on the experimentation with a virtual coach system on users while cycling on a stationary bike. The virtual coach, developed by the Philips Research Centre, is projected on a screen surrounded by a naturalistic landscape. Their results show a good users reaction to the stimuli provided by the virtual coach with some other interesting results about the effectiveness of informations provided by the coach during the workouts.

Your Shape Fitness Evolved [12] is a fitness game designed for Microsoft Kinect. The software guides users step by step during indoor workouts and let them customize their workouts in terms of goals to reach. The game allows users both to keep track/share their statistics through the community and to challenge other users.

Batussi et al. [5] developed a PocketPC application called Mobile Personal Trainer (MOPET) aimed to support users during their workouts. MOPET uses GPS device and vocal cues during training sessions. The application makes use of an embodied virtual trainer that guides users showing the proper execution of exercises.

Toscos et al. [13] propose Chick clique that aims to push teenage girls to adopt a correct lifestyle. The application collects informations about the caloric content of popular foods and the amount of steps necessary to burn them. The software promotes social interactions by means of SMS to boost a friendly competition among users.

Consolvo et al. [14] have developed a mobile application called Huston that, through a pedometer, counts and records the number of steps done. The results collected during Huston's experimentation are very interesting. Authors derived four key design requirements that may help developing such applications:

- Users want accurate measurements of their activities
- Deep use of long-term statistical reports
- Strong social interaction support
- Applications must be developed taking into account users lifestyle habits

Oliveira et al. [15] presented TripleBeat a mobile phone application that makes use of both an ECG and an ac-

 ISBN: 978-1-61208-228-8

celerometer to push runners to achieve their goals in terms of a certain heart rate. The results of the experimentation have revealed the importance of a well-designed and intuitive graphical interface to improve self-awareness and the effectiveness of virtual competition to enforce users motivations.

Mulas et al. [17] propose an Android-based mobile application called Everywhere Run! that aims at motivating and supporting people during their running activities. The application behaves as a virtual personal trainer helping runners to keep a predefined pace. With respect to other similar applications, Everywhere Run! is designed to promote interactions between users and real coaches by means of a community of runners.

Nike+GPS [16] as been designed by Nike for the IOS operating system. It is one of the most complete and popular applications in sport and health area. Some of the most important strengths are:

- Intensive use of social networks
- A well-designed system for the management of vocal cues and music
- A dynamic web community where users can create their workouts and at the same time interact with other sportsmen

There are a number of applications similar to these ones. Just to name a few (see Section "References" for websites): Runtastic, Endomondo, RunKeeper, MapMyFitness, Adidas miCoach and so on. All these products provide more or less the same features:

- Route and workout data tracking
- Workout statistics
- Results sharing through social networks

The last application we will discuss is the most similar to our proposal; it is called Softrace [23]. In this application, in addition to the just discussed features, users can compete with other people in a sort of real-time competition.
The trait that sets us apart from this proposal is the totally different implemented concept of real-time race. In Softrace a user can start his race when he wants. He then races against other people that are already running at that moment. Thus, there is not a real concept of "race". Our application instead implements a race in its classical meaning, thus with both a priori known start time and distance, both the same for all the participants. Furthermore, in Softrace, there is not the concept of virtual events associated with real ones.

III. Everywhere Race!

Everywhere Race! is an Android application (version 1.6 and newer compliant) designed to motivate people to exercise regularly. It is already available in the Google Play Store [24]. It engages people through the totally new concept of real-time virtual competition deeply based on fun and social interactions. The software, for the first time ever, makes it possible to create a real-time race in the classical sense of the word. Users can choose almost any kind of speed-based sports and challenge their friends and other people from all over the world seamlessly from the application.

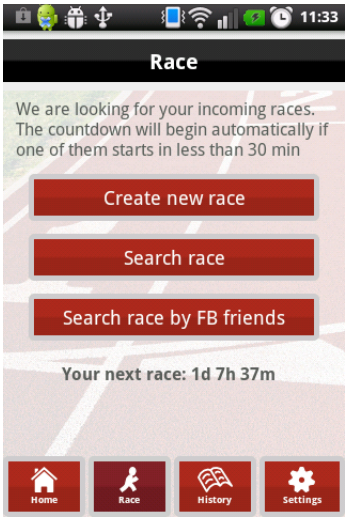

Figure 1. Main menu

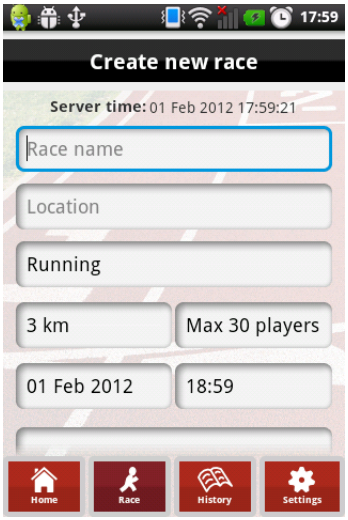

Figure 2. Create new race menu

It follows a more detailed description of *Everywhere Race!*. Figure 1 shows the main race menu. From here the user can create a new race, perform a search for existing ones based on common attributes (e.g., sport, distance, starting time and so on) or directly search for races in which his friends are involved in. As it is possible to see from the screenshot, the application automatically displays the remaining time to the upcoming user's race. In this example, the next user's race will be in a little more than a day.

From the "Create new race" menu (see Figure 2), it is possible to create a new race by choosing the wanted race

 ISBN: 978-1-61208-228-8

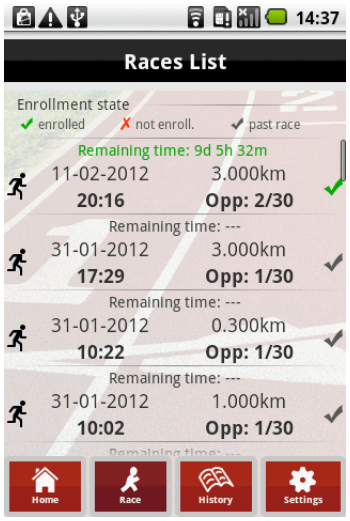

Figure 3. Search results

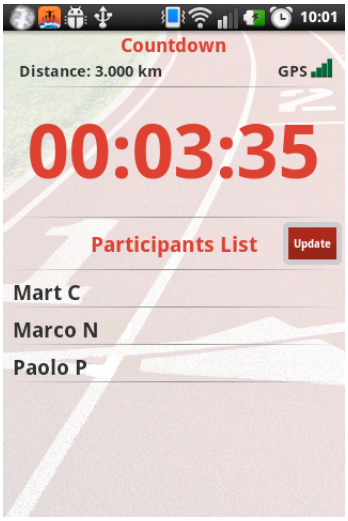

Figure 4. Countdown dialog

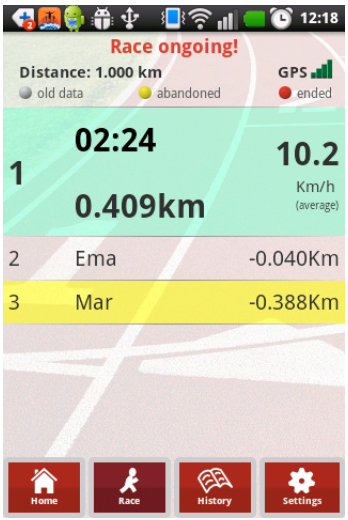

Figure 5. Ongoing race

parameters such as, for example: name, length, number of opponents and start time. As stated above, the application allows two kinds of searches that return a list of races matching the query parameters. An example of search result is shown in Figure 3.

The user's membership state, the start time, the distance and the number of opponents enrolled so far are shown for every returned race. It is possible to have more details by tapping on the row corresponding to a race. We have just described the steps needed to create/find an existing race in order to enroll on it.

Let us speak now about the funniest and more important feature provided by the application, the race. As clearly stated in the top of the screen in the Figure 1, a countdown will automatically start if the next user's race begins in less than thirty minutes. Figure 4 shows that case, where a race will start in less than four minutes.

Note: to avoid misunderstandings, till the end of this section, we will use the word "user" for indicating the user of the current smartphone and the word "opponents" for indicating the other remote users (his competitors).

Figure 5 shows the user screen during a race. The most important information needed while racing is reported on this screen. Starting from the top of the screen, the user can see the race nominal distance and some other data like the GPS signal strength and the meaning of the colours that may appear while racing to report the state of the opponents. The rest of the screen shows a snapshot of the race at a certain point in time. In this particular case (see Figure 5) it is reported the state of the race among three contenders at about two minutes from start. The first row reports data of the current first player. In this case the first player corresponds to the current user and this explains why the row is bigger than others. The cell reports the user's position, race time, distance and the average speed. The other two rows show data of the current second and third player with their gap (expressed as a distance) from the user, respectively.

As soon as the current user covers the whole race distance, the race for him is finished, as in a real one. At this point, the user can choose whether to publish his race result in his Facebook wall or not. Finally, EWR shows the actual classification (see Figure 6). Note that other opponents can be still racing (again, as in a real event), thus the classification can be not final yet.

IV. Experimental Results

To evaluate the effectiveness of the proposed application we submitted a subjective evaluation test to a sample of 35 volunteers that used *Everywhere Race!* for 30 days.

 ISBN: 978-1-61208-228-8

Figure 6. Race results screen

The sample was composed of 25 male users and 10 female users aged between 19 and 40. Ten users regularly practiced sport at amateur level (6 males, 4 females), whereas the remaining testers were mostly sedentary. Among active users 4 exercised about four times a week, while the others from two to one times a week for an average training session duration of 30 minutes. Twenty-seven users (5 females and 22 males) were non-technical and they had never used before any application as a support to physical activity whereas the remaining part of testers already had some experience with such kind of applications.

With these evaluation tests, we wanted to investigate the influence of our application on users motivation in order to both validate current application features and understand how to proceed for future developments.

There are several definitions of motivation:

> the psychological process that gives behavior purpose and direction (Kreitner)
>
> a predisposition to behave in a purposive manner to achieve specific, unmet needs (Butford, Bedeian and Lindner)
>
> an internal drive to satisfy an unsatisfied need (Higgins)
>
> the will to achieve (Bedeian)

Ryan et al. [25] propose the following subclassification of motivation:

- Intrinsic motivation: refers to doing something because it is inherently interesting or enjoyable
- Extrinsic motivation: refers to doing something because it leads to a separable outcome

Vallerand et al. [26] studied the differences between intrinsic and extrinsic motivational factors in sports. They claim that intrinsic motives are generally considered to be more relevant than extrinsic ones especially for non-habitual sportsmen. These results are generally considered valid, although other studies (see [27]) show that the user's gender may affect the predominant motive. Females appear to have more intrinsic motives than males which, vice-versa, have more extrinsic motives than females.

The description of how *Everywhere Race!* impacts on motivational aspects has been evaluated follows. For this purpose, we chose to adopt the Exercise Motivations Inventory - 2 (EMI-2) developed by Markland et al. [28]. EMI-2 is composed of 51 items belonging to 14 scales. Testers are asked to rate each item on a five-point scale ranging from 0 ("not at all true for me") to 5 ("very true for me").

Scale scores are obtained by calculating means for each item belonging to the appropriate scale. The graphic in Figure 7 shows the obtained results.

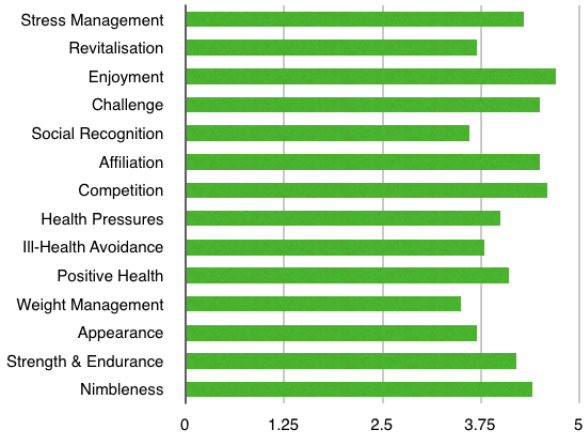

Figure 7. EMI-2 results

All the scales received good scores meaning that the application is a valid mean to help people to start working out. Some of the most important scales that we wanted to test were those related to enjoyment, challenge, affiliation and competition. This is because we aimed at evaluating the innovative features provided by *Everywhere Race!* in order to understand if our application is evolving in the right direction.

We submitted an additional questionnaire to deeper investigate the effects of the application on users sport habits. Testers were asked to rate each question with the same scale (ranging from 0 to 5) used to rate the EMI-2 items.
Here are the questions:

1) "Did EWR help you to improve performances?"
2) "Were social features important to improve your performances?"
3) "Did EWR change your sport habits?"
4) "Will you continue to use EWR in the future?"

The average results shown in Figure 8 put in evidence that the majority of users perceive the application as a valid

 ISBN: 978-1-61208-228-8

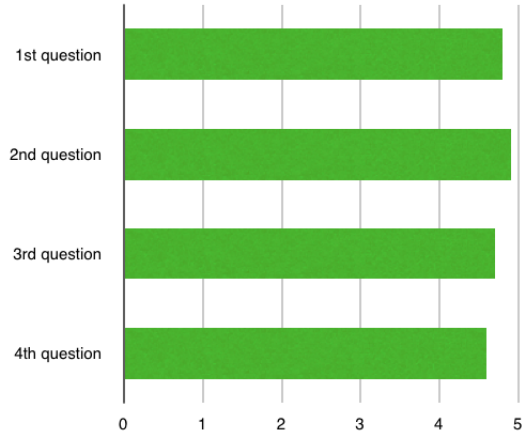

Figure 8. Questionnaire results

tool that helps to achieve sport goals in a more enjoyable and regular manner.

As it is possible to see, despite the limited sample of test users, we have obtained encouraging preliminary results. The positive trend emerging from our tests shows that the application may help to increase motivational factors through this new engaging and social way of active gaming.

V. Conclusion and Future Work

Nowadays, many studies have been carried out trying to find possible solutions to the alarming increase of health diseases. These pathologies are mainly related to an unhealthy lifestyle often due to a lack of physical activity.

In this paper, we presented *Everywhere Race!*, a smartphone application that aims at attracting as many people as possible to start working out in a totally engaging and social fashion. For the first time ever, users can challenge people from all over the world in real-time races very similar to real ones directly from a smartphone. It is like a game, but the participants can move for real! All participants start the race at the same time wherever they are. At any moment in time, all players can see both their position in the race and that of their opponents. At the end of the race, *Everywhere Race!* shows the final classification with arriving positions, times and speeds as in a real race.

As proven by our tests, one of the strengths of the application is its fun and social-oriented design. These features allowed us to exploit the complex social dynamics that has been proved to be very important and effective for people engagement, especially in sports (see, for example, [29][30]).

To evaluate the effectiveness and ease of use of *Everywhere Race!*, we have conducted some preliminary tests with a group of 35 volunteers. The tests have been designed to actually prove the benefits of our software to users motivation. Our results showed that the majority of people that in the past had a sedentary lifestyle got motivated to working out more often. All these benefits, as shown by numbers, have a close relationship with the innovative features of the application.

Our research will proceed in several directions. First, we are currently developing a web community associated with the application to enforce social interaction among users. Second, we plan to add new features to the application. Among the most forthcoming: more detailed users and community statistics, more social functions and speech synthesis support. Thanks to these improvements we aim to gain a greater visibility on Google Play Store in order to perform more accurate tests on a more statistically meaningful sample of users.

References

[1] Android. (2012). [Online]. Available: http://www.android.com

[2] D. Warburton, C. Nicol, and S. Bredin, "Health benefits of physical activity: the evidence," *CMAJ*, vol. 174, no. 6, mar 2006.

[3] R. Pate, M. Pratt, and S. B. et. al., "Physical activity and public health. a recommendation from the centers for disease control and prevention and the american college of sports medicine," *JAMA*, vol. 7, 1995.

[4] WHO-ObesityAndOverweight. (2012, May). [Online]. Available: http://www.who.int/mediacentre/factsheets/fs311/en/

[5] F. Buttussi, L. Chittaro, and D. Nadalutti, "Bringing mobile guides and fitness activities together: a solution based on an embodied virtual trainer," *MobileHCI 06: Proceedings of the eighth conference on human computer interaction with mobile devices and services, ACM Press*, pp. 29–36, 2006.

[6] Facebook. (2012). [Online]. Available: http://www.facebook.com/

[7] S. Senevirathne, M. Kodagoda, V. Kadle, S. Haanke, T. Senior, and B. Heller, "Application of serious games to sport, health and exercise," *Proceedings of the 6th SLIIT Research Symposium*, 2011.

[8] J. Sinclair, P. Hingston, and M. Masek, "Considerations for the design of exergames," *in Proc. 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*, pp. 289–295, 2007.

[9] J. Hoysniemi, "International survey on the dance dance revolution game," *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment*, vol. 4, Apr 2006.

[10] C. Jayant and T. S. Saponas, "Mariofit: Exercise through mobile entertainment," 2005.

[11] W. IJsselsteijn, Y. de Kort, J. Westerink, M. de Jager, and R. Bonants, "Fun and sports: Enhancing the home fitness experience," *In ICEC 2004: Proceedings of the 3rd International Conference on Entertainment Computing, Springer*, pp. 46–56, 2004.

 ISBN: 978-1-61208-228-8

[12] YourShapeFitnessEvolved. (2012). [Online]. Available: http://your-shape-fitness-evolved.ubi.com/2012/

[13] T. Toscos, A. Faber, S. An, and M. P. Gandhi, "Chick clique: persuasive technology to motivate teenage girls to exercise," *CHI '06 extended abstracts on Human factors in computing systems, Montral, Qubec, Canada*, April 2006.

[14] S. Consolvo, K. Everitt, I. Smith, and J. A. Landay, "Design requirements for technologies that encourage physical activity," *Proceedings of the SIGCHI conference on Human Factors in computing systems, Montral, Qubec, Canada*, April 2006.

[15] R. de Oliveira and N. Oliver, "Triplebeat: Enhancing exercise performance with persuasion," *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services, Amsterdam, The Netherlands*, September 2008.

[16] Nike+GPS. (2012). [Online]. Available: http://nikerunning.nike.com

[17] F. Mulas, S. Carta, P. Pilloni, and M. Manca, "Everywhere run: a virtual personal trainer for supporting people in their running activity," *Proceeding ACE '11 Proceedings of the 8th International Conference on Advances in Computer Entertainment technology, ACM Press*, Nov 2011.

[18] Runtustic. (2012). [Online]. Available: http://www.runtastic.com/

[19] Endomondo. (2012). [Online]. Available: http://www.endomondo.com/

[20] RunKeeper. (2012). [Online]. Available: http://runkeeper.com/

[21] MapMyFitness. (2012). [Online]. Available: http://www.mapmyfitness.com/

[22] miCoach. (2012). [Online]. Available: http://www.adidas.com/fi/micoach/

[23] SoftRace. (2012). [Online]. Available: http://www.softrace.net/

[24] EverywhereRace! (2012). [Online]. Available: http://www.everywhererace.com/

[25] R. Ryan and E. Deci, "Intrinsic and extrinsic motivations: Classic denitions and new directions," *Contemporary Educational Psychology*, no. 25, pp. 54–67, Jan 2000.

[26] R. Vallerand, E. Deci, and R. Ryan, "Intrinsic motivation in sport," *Exercise & Sport Sciences Reviews*, vol. 15, pp. 389–426, Jan 1987.

[27] M. Fortier, R. Vallerand, N. Brire, and P. Provencher, "Competitive and recreational sport structures and gender: A test of their relationship with sport motivation," *International Journal of Psychology*, no. 26, pp. 24–39, 1995.

[28] D. Markland and D. Ingledew, "The measurement of exercise motives: Factorial validity and invariance across gender of a revised exercise motivations inventory," *British Journal of Health Psychology*, no. 2, pp. 361–376, 1997.

[29] B. Schellenberger, "The significance of social relations in sport activity," *International Review for the Sociology of Sport*, vol. 16, pp. 69–77, Jan 1981.

[30] A. Sato, A. Wilkinson, and L. Koehly, "Social influence and motivation to change health behaviors among mexican-origin adults: Implications for diet and physical activity." *American Journal of Health Promotion*, vol. 26, no. 3, pp. 176–179, Jan 2012.

[31] A. Wroblewski, F. Amati, M. Smiley, B. Goodpaster, and V. Wright, "Chronic exercise preserves lean muscle mass in masters athletes," *The Physician and Sportsmedicine*, vol. 39, no. 3, Sept 2011.

 ISBN: 978-1-61208-228-8

Fora for Distance Education: Another Way for Analysis of Discussions

Kiriakos Patriarcheas and Michalis Xenos
School of Sciences and Technology, Computer Science
Hellenic Open University
Patras, Greece
{k.patriac, xenos}@eap.gr

Abstract— **This paper refers to fora for distance education. The present research focuses on content analysis, a technique frequently used to approach issues concerning asynchronous computer mediated discussion groups. There is a variety of approaches, varying on the level of detail and on the type of categories of analysis they use. The content category is presented as a unit of analysis for the decodification of messages in the asynchronous distance education fora in which the modelling is incorporated in a formal language and the development of a respective system created by Hellenic Open University for this purpose. The creation of this system makes an important contribution to the decoding of discussions in fora, and aims at summary identification of discussions which do not develop in the desired way.**

Keywords-E-Learning; Asynchronous distance education fora; Content analysis; Formal language; Modelling; AdaBoost; Naive Bayes; 1-Nearest Neighnor; WINNOW.

I. INTRODUCTION

During the last decade, an increasing number of educational institutions, as well as companies, apply asynchronous educational services via internet [1][2][3]. One of the means utilised in distance education during the last decade is the electronic fora (fora hereinafter). Research efforts on distance education fora, at an international level, began during the ’90s [4][5][6]. However, this is a field, which requires constant updating and redefinition. Given, also, the fact that the practice of distance education during the last decade has acquired new features, both in its methodology and in the tools which are utilised, the further exploration of this field becomes necessary.

The structure of this article is as follows: the theoretical framework section is a short description of the respective assignment on the content analysis technique of asynchronous discussions at distance education fora. The role of the fora of Hellenic Open University (HOU hereinafter) concerning the educational procedure is described. The unit of analysis, which was used, is then presented, followed by the integration of the message content category as a unit of analysis in formal language. A presentation of the system of automatic text classification and the association with the message content category follows, including the description of the results of the experiments performed to control system operation. The paper closes with a discussion on the necessity of this system and the conclusions of this article.

II. THEORETICAL FRAMEWORK

Although researchers seem to agree that collaboration may encourage the learning procedure [1], there is no clear theory available to guide research on computer mediated interaction [2], empirical markers which shall be the base of a codification tool as a standard against which to evaluate whether or not effective learning is occurring through the online discussions [3]. In the last few years, numerous efforts to approach this issue were made, stemming from different theoretical backgrounds. Indicatively, Henri [4] uses the point of Cognitive and metacognitive knowledge, while others [5][6] the point of Critical thinking.

As is shown from all the above, an important issue arising is the unit of analysis which shall be used for the content analysis. Fahy et al. [7] consider each single sentence as one unit of analysis, and Pena-Shaff and Nicholls [8] uses the sentence as unit of analysis, trying to approach it at a paragraph level. Others choose the definition thematic unit (or otherwise of a “theme” or an “idea”) to be their unit of analysis [4][5][9]. Another approach [3][6][7][10] is to consider the whole message that a student enters at a specific moment in the conversation as the unit of analysis. Jarvela and Hakkinen [11] choose a Complete discussion, while during the last years there has been an approach of multiple point both at a micro and at a macro level [12]. Further down, a comprehensive review is presented in a table form (Table I), referring to the unit of analysis used by this field’s researchers.

TABLE I. OVERVIEW OF THE CONTENT ANALYSIS SCHEMES

Instrument	Theoretical background	Unit of analysis
Henri (1992)	Cognitive and metacognitive knowledge	Thematic unit
Newman et al. (1995)	Critical thinking	Thematic unit
Zhu (1996)	Theories of cognitive and constructive learning – knowledge construction	Message
Gunawardena et al. (1997)	Social constructivism – knowledge construction	Message
Bullen (1997)	Critical thinking	Message
Fahy et al. (2000)	Social network theory – Interactional exchange patterns	Sentence
Veerman & Veldhuis-Diermanse (2001)	Social constructivism – knowledge construction	Message
Jarvela & Hakkinen (2002)	Social constructivism – perspective taking	Complete discussion
Lockhorst et al. (2003)	Social constructivism – learning strategies	Thematic unit
Pena-Shaff & Nicholls (2004)	Social constructivism – knowledge construction	Paragraph
Weinberger & Fischer (2006)	Social constructivism –knowledge construction	Micro and macro-level

 ISBN: 978-1-61208-228-8

III. THE CASE OF HELLENIC OPEN UNIVERSITY

HOU is the eminent educational institution offering distance education in Greece. Today, HOU has 30,557 students (17,889 undergraduate, 12,600 postgraduate and 68 PhD candidates); it is staffed by 1642 professors (only 42 of which are permanent and the rest are associate professors-counsellors).

The HOU's structural educational unit is the course module; presently, 203 course modules are offered by HOU. An important supportive mean of the educational procedure is the fora of HOU, which contribute both to the organization of the studies during the course module as well as to the elaboration and development of what the student have already studied.

The HOU's fora offer important help during the educational procedure. They may also contribute to the following:

a) as for the organization of the studies during the course module:

- to the communication between the tutor and the students (regularity of contacts, subject, resolution of "technical" problems etc.).
- to the organization of homework (method of use of the teaching material and the preparation of the activities, exploitation of the literature and the other sources, timetables, encountering problems related to it etc.)
- to the supply of information about the advisory meetings (their number, their duration, the timetables, the goals, their content and methodology applied, problems' encountering as for the ability to attend them etc.).
- to supply clarifications about the procedure of preparation and evaluation of the written assignments (form, method of preparation, evaluation criteria, ways to be supported by the tutor etc.).
- to inform about the procedure of final exams.

b) as for the elaboration and development of what the students have already studied, the HOU's fora may be exploited for:

- presentation of consolidation exercises, short suggestions, presentation of examples, methodologies, literature etc.,
- resolution of questions and the supply of clarifications about the teaching material.
- interconnection between what is already studied and the next chapters and the following written assignment.

In the discussion threads of each course module, the tutor and all the students of the course module have the chance to participate. As for the students of informatics, for the 16 course modules of informatics (for undergraduate level) offered by HOU, by the time this research was conducted, there were 753 discussion threads created with 6,663 messages. Concerning the evolution of the HOU forum's use, indicatively, at the course module "Introduction to Informatics" (INF10), during the last three academic years there is a great increase in the number of messages: 1808(2009-10), 1942 (2010-11) noted and 2913 (2011-12).

Given the big flow of information transferred through fora of HOU, simulated the development of a formal language to interpret messages in the fora of HOU, a system based on modelling with the use of a formal language was created, entering threads from discussion fora and exporting the respective strings in an automatic way.

IV. THE UNIT OF ANALYSIS

Given that the choice of a unit of analysis is dependent on the context and should be well-considered, because changes to the size of this unit will affect coding decisions and comparability of outcome between different models [14], as well as given the fact that Schrire [15] refers to a dynamic approach in which data is coded more than once and the grain size of the unit of analysis is set, depending on the purpose and the research question, it was decided not to take into consideration the discussion thread, not even the message as unit of analysis, nor the paragraph or the single sentence.

It was decided to use as unit of analysis, the category of the message's content, as for the observation of the discussion threads, it was noticed that there are cases of messages which may comprise two (or/and more) content categories, e.g., a question about the next advisory meeting and a reply to a question concerning the study of the educational material.

Thus, in that case, the analysis at a message's level used by some researchers [3][6][10][13] is insufficient for the exploitation of information that shall arise aiming to reach educational conclusions, as it is obvious that in a message more content categories may coexist.

Furthermore, the analysis at a level of a single sentence used by some researchers [7] could not be taken as a single unit of analysis since a content category may extent to two or/ and more sentences. For the same reason the unit of analysis at a paragraph level, used by Pena-Shaff and Nicholls [8], was not chosen. In addition, in our case, not even the Complete discussion used by Jarvela and Hakkinen [11] as unit of analysis cannot be exploited, since it is noted that in a discussion thread there may exist many more than one content categories. Otherwise, the title of a discussion thread may not be representative of this and the discussion may extent to more than one subjects.

According to the study of the messages of INF10 for academic years 2009-2012 the messages as for their content may concern (in brackets you see the respective symbols used in formal Language): study of educational material (*M*), questions/answers for exercises – assignments (*X),* presentation of sample assignments by tutors (*P*), instructions (*I*), assignment comments, corrections (F), student comments on assignments (*D*), sending – receiving assignments (*J*), sending - receiving grade marks (*G*), notification of advisory meeting (*V*), and pointless message (*L*).

 ISBN: 978-1-61208-228-8

V. INTEGRATION OF THE MESSAGE CONTEXT CATEGORY AS A UNIT OF ANALYSIS IN FORMAL LANGUAGE

According to the study of the messages on HOU's fora during the academic years, a Language was developed, which is defined by mathematic terms and represents the messages using as unit of analysis the category of the message content. More specifically:

There are two categories of communication's carriers: a) Tutors, b) Students

For brevity reasons, tutors shall be symbolized with *T* and students with *E*

As for the type of message, they are discerned to questions and replies (symbols *q* and *a* respectively).

As for their content category, we have the symbols aforementioned in the previous section: *M, X, P, I, F, D, J, G, V*.

Finally, the order in which the above symbols appear is: a) the message carrier, b) the type of message and c) the content category to which the message belongs.

Thus, the Language contains:

a) Terminal symbols alphabet V_T, where V_T = {T, E, q, a, n, M, X, P, I, F, D, J, G, V, L }

b) Non terminals alphabet V_N, where V_N = {u, r, y, c}, more specifically :

r: represents the message carrier (where *T* for tutors and *E* for students)

u: represents a pair *yc* i.e., a message type *y* (whether it is a question *q* or an answer *a*) followed by its content category.

c) The grammar P

A set of rules of the form α → β, where α and β sequences containing terminal and non-terminal symbols and α is not an empty sequence, as follows:

1. S → ruS	8. y → q	15. c → F
2. S → ε	9. y → a	16. c → D
3. u → uyc	10. y → ε	17. c → J
4. u → ε	11. c → M	18. c → G
5. r → T	12. c → X	19. c → V
6. r → E	13. c → P	20. c → L
7. r → ε	14. c → I	21. c → ε

where ε stands for an empty symbol

d) Symbol S where every sentence generated starts with this symbol.

According to the above, when a message should be represented concerning a student's message, addressing a question about the study of the educational material, followed by another student's question about the following assignment and at the end of the thread there is the reply of the tutor both for the study of the material and for the following assignment, it shall be represented as follows: *EqMEqXTaMX* (*E* for the student's capacity, *q* for the question, *M* as it concerns the study of the educational material, *X* for the fact that the next message concerned an assignment, *T* for the tutor's capacity, *a* for the fact that it is an answer, *M* for the fact that this reply concerns the study of educational material and *X* for the fact that the second part of the message concerns an assignment.). According to the above, the sequence *EqMEqXTaMX* constitutes a sentence of the *Language* because:

Rule: (1) (1) (1) (3)
S —> ruS —> ruruS —> rururuS —> ruycruycruycS

(4)(6)(8)(11) (4)(6)(8)(11)
—————> EqMruycruycS —————> EqMEqXruycS

(3) (2)(4)(5)(9)(10)(12)
—>EqMEqXruycycS ——————————> EqMEqXTaMX

As it is obvious from the example, while to the first two messages corresponds one content category *M* and *X* respectively, at the 3rd message there are two content categories *MX*.

VI. THE SYSTEM - TIME ASSOCIATION

According to this approach, a system of automatic classification was developed, which comprised the following:

a) Data filtering: where some web pages are considered as input accommodating the discussion threads of a distance education forum of HOU (which include a great deal of irrelevant data containing essential information concerning the educational procedure, e.g., titles, images etc.) and creates a temporary file with the "useful" part (User name, date, message's content), which may become a source of information for educational conclusions.

b) Storage of root files: which is a dynamic way according to which word or phrases or symbols roots are stored, as well as the respective terminal symbols q if it is a question or a if it is an answer. The same was also done for the storage of information necessary in the determination of content category of a message, i.e., if it is about study, assignment, comment etc. or combination of them (e.g., a message concerning both the study and an assignment). To wit, it takes as input couples of information of the type root of a word or phrase and terminal symbol of the content category (M, X, P, I, F, D, J, G, V, L). As it is obvious, the system provides the ability to add further content categories if necessary.

c) Strings' production: receiving as input the temporary file with the "useful" information (User name, date, message's content) and the files with the couples of roots words/ phrases/ symbols and terminal symbols and presents the respective strings with the relative extensible file, so as the results to be kept for further exploitation (Figure 1).

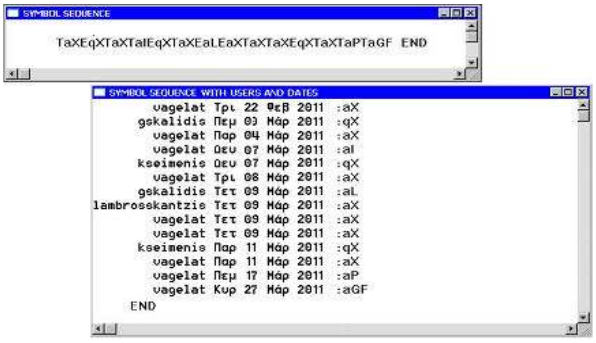

Figure 1. Representation of a discussion thread both in simple string and also after the addition of User names and dates (Days: Κυρ=Sun, Δευ=Mon, Τρι=Tue, Τετ=Wed, Πεμ=Thu, Παρ=Fri, Σαβ=Sat, Months:Φεβ=Feb, Μάρ=Mar, Απρ=Apr).

 ISBN: 978-1-61208-228-8

It is worthy to note here that this specific system incorporates the sense of time along with its association with each of the nine (9) categories of message content (except from the pointless message category) chosen as unit of analysis. More specifically and given that within a message (as it is deduced both from literary review and from the observation of the fora of HOU) more than one contents may exist, the dates are recorded for each such case and not simply in each message.

In fact, after each couple yc, there is a date's record. Certainly, so as to effectuate the above procedure nine (9) stacks were used – as many at the message's content categories, each one having as many figures as the number of appearance of terminal symbols resulting from the non-terminal symbol c. Consequently, time differences may automatically exist (in days, if from each current date, by content category, it is deduced the previous one) and thus there may arise another nine (9) respective stacks with the above date references. Of course, the length of these stacks is equal to the length of dates minus one (-1), i.e., apart from the initial message, which is considered to be the point zero (0), where the numbering of the time differences begins. The contents of the stacks of time differences may constitute an important criterion, which may participate as such (in combination with other criteria) in case of evaluation of a forum's consequences to the educational process.

Also, the final form that will have the 9 stacks with the dates they will be as follows (Table II):

TABLE II. DATE'S STACKS

null										
3/20/12	4/1/12	4/2/12	4/5/12	4/6/12	4/7/12	4/7/12	4/7/12	4/9/12	4/9/12	null
4/15/12	null									
4/5/12	null									
4/25/12	null									
null										
null										
4/25/12	4/28/12	null								
null										

while the counter table (Table III) has the following contents:

TABLE III. COUNTER TABLE CONTENTS

1	2	3	4	5	6	7	8	9
0	10	1	1	1	0	0	2	0

As for the time differences stacks (Table IV) has the following form:

TABLE IV. TIME DIFFERENCES STACKS

Content	Appearances	1	2	3	4	5	6	7	8	9
M	*0*									
X	*10*	9	1	3	1	1	0	0	2	0
P	*1*	0								
I	*1*	0								
F	*1*	0								
D	*0*									
J	*0*									
G	*2*	3								
V	*0*									

VII. EXPERIMENTS

A. Phase A

At first, experiments were carried out by using 80 discussion threads of the INF10 module of the academic year 2010-11. Given that 219 threads with 1,942 messages had been created throughout the year, there was the ratio of approximately 9 messages (in particular 8.87) per discussion thread. Therefore, out of the 80 selected threads, an effort was made to use those containing 8 or 9 messages for the purpose of experimental control. Thus, we finally chose 80 discussion threads with 712 messages in total (average 8.90 messages/thread).

At the first experimental operation, the word root files in relation to both the type (question/answer) and (mainly) the content category of message contained 18 and 92 entries respectively. Under these conditions (Table VI), we ended up having 58 discussion threads with no errors and 16 threads with only one wrong symbol (compared to what was expected). Namely, out of (approximately) 9 messages (of each of the 16 threads), 8 of them were correct and one message was wrong because it did not contain not even one of the 92 provided word roots. Respectively, there were 5 threads with two errors and 1 thread with more errors (this thread was created before Christmas holidays and its messages contained mainly wishes). We should note here that there has been no error regarding the type of messages (question/answer), only in terms of determining the content category.

Following the observation/study of messages in the 21 threads that contained 1 or 2 errors, 49 additional word roots (concerning the content category) were recorded and it was decided to enter them in the root file. The experimental operation performed in the same 80 threads had, clearly, better results, with total success in 70 threads, one wrong symbol in 8 threads, two errors in 1 thread, and 1 thread that did not actually refer to educational content (Table V).

At this point, it should be clarified that the control of the results produced by the system in this phase (A), was conducted with manual comparison of all the messages in the discussion threads that were used in order to control system reliability at the first degree.

TABLE V. EXPERIMENTAL OPERATION - PHASE A'

2010-11	**1st Exp. operation**	**2nd Exp. operation**
Threads	80	80
Messages	712	712
Messages/Thread	8.9	8.9
Full success (threads with no errors)	58	70
Threads with one error	16	8
Threads with two errors	5	1
Threads with more errors	1	1
Correct messages interpretation	677	693
Wrong messages interpretation	35	19

B. Phase B

Given that the 8 discussion threads with one error were found not to have any common word root feature that would adequately correspond, we decided to initiate the second

 ISBN: 978-1-61208-228-8

experimental phase (B'). Classification was performed according to international literature [16-21], using the algorithms indicated for this purpose: Naive Bayes (NB), 1-Nearest Neighnor (1-NN), WINNOW and discrete AdaBoost (in the form generalized by Nock and Nielsen [19] based on Freund and Schapire [20]).

During this phase every algorithm was formed using the data collected from the academic year 2010-11. Subsequently, a group of data for two other academic years (2009-10 and 2011-12) was also collected. The results show that the discrete AdaBoost algorithm produced the greatest accuracy. This result complies with Bloehdorn and Hotho [21] who used the discrete AdaBoost algorithm in a similar experiment. The accuracy is denoted in the Table VI.

TABLE VI. ACCURACY OF ALGORITHMS FOR THE ACADEMIC YEARS 2009-12

	2009-10	2010-11	2011-12	Average Accuracy (1)	Average accuracy (2)
In thread level					
AdaBoost	75.11	80.08	87.21	80.64	80.80
Naive Bayes	72.47 *	77.83	86.18	78.66	78.82
1-Nearest Neighnor	73.45	76.66	83.65*	77.77	77.92
WINNOW	70.13 *	73.24*	83.10*	75.34	75.49
In message level					
AdaBoost	92.36	95.19	97.89	94.96	95.15
Naive Bayes	89.11*	92.51	96.73	92.59	92.78
1-Nearest Neighnor	90.31	91.13	93.89*	91.60	91.78
WINNOW	86.23*	87.06*	93.27*	88.67	88.85

The star (*) indicates that the algorithm which were used ,performed statistically better than the specific classifier according to t-test with $p<0.05$. In all the other cases, there is no significant statistical difference between the results (Draws).

The average accuracy (1) corresponds to the total number of threads and messages, while in (2) the years have an equal participation (1/3) in the total average.

It is worth noting that in all cases, the type of message has been correctly identified (i.e., whether it is a question or answer), and therefore any errors concerned the content category (as noticed in the initial experimental operation). Results are shown in Table VII.

TABLE VII. RESULTS FOR THE ACADEMIC YEARS 2009-12

Year	2009-10	2010-11	2011-12	Total
Threads	356	219	178	753
Messages	1808	1942	2913	6663
Messages/Thread	5.08	8.87	16.36	8.85
Threads with no	269	191	159	707
Percentage	75.56%	87.21%	89.33%	80.64%
Correct messages	1678	1901	2864	6443
Percentage	92.81%	97.89%	98.32%	96.70%

The above results, which followed a calibration process of repeated readjustment, were deemed satisfactory (98.32% correct message interpretation for 2011-12) and in the end, the development of this system gives a clear affirmative answer to the question "is there an automated method to interpret messages in a distance education forum?" Therefore, by using this system, it is now possible to read study and classify, within a few minutes, a large number of messages (6,663 messages) which took 12 months to be completed within the framework of this paper.

VIII. DISCUSSION

At this point, it is important to initially discuss the need for such a system using the content category as analysis unit. As it is deduced from the related literature review, it is concluded a gap as for the methodological approach which is to be based on modelling with the use of formal language and which will examine the content analysis from the message content category view. This research aims to contribute to the covering of this gap, consequently a need arose to create a system to interpret the discussion of a forum about distance education in a structured way, through the approach presented above and at the same time taking into consideration the concept of time and producing respective results, so as to help the participating students to improve their educational practices.

In fact, this system defines a "code" clarifying some issues which determine both the quality of the communication relations and the educational principles of teaching/ learning procedure. In other words, this system aims to encourage towards the direction of the use of "good" or "desirable" educational techniques, adding up to the distance education. At this point it should be clarified that the development of this system does not intend to disorientate from the basic principles of distance education, but aims to contribute to its further development and upgrading and to act auxiliary and not in excess. The results of its use shall constitute data for the creation of a database [22] aiming to investigate the effects of fora in educational procedure from the point of causal interpretation point of view. Given that the HOU is not a conventional university (with the features of a homogenous student community), but it addresses to adults with special educational needs and incongruity (both as far as their age, their professional and family obligations are concerned), the future research access to such issues becomes particularly important.

More specifically, the system in the future shall collect the students' particulars (e.g., marital status, age, sex, profession etc.), their performance at course modules of HOU (final mark, assignments' marks, effort of success of the course module) and the strings produces by the system so as to interpret the messages of HOU's fora, in order to reach educational conclusions in combination with the use of the a tool, weka type. In fact, this application refers to a wider field of interdisciplinary encounter, by the merge of cognitive theories and artificial intelligence.

IX. CONCLUSIONS AND FUTURE WORK

The practice of distance education during the last years has acquired new features, both in relation to methodology and in the tools it uses. It is also a fact that the subject of electronic fora in distance education is a dynamically formed field requiring constant updating and redefinition. A big part of the research presented in the international literature concerning distance education's fora, refer to the content analysis, which principally aims despite the fact that this research technique is frequently used, though there are

 ISBN: 978-1-61208-228-8

still no standards established. There is a variety of approaches, varying both at detail's level and at the type of categories of analysis they use. As it was deduced from the above presentation and study of the discussion threads of HOU, it was noticed that there are cases of messages, which may comprise two (or/and more) content categories, e.g., a question about the next advisory meeting and a reply to a question concerning the study of the educational material. For this reason, this paper uses the content category as unit of analysis for the messages' interpretation in Asynchronous distance education fora and for this purpose incorporates it in a modelling in a formal language. Furthermore, time indexes of participation were integrated in combination with the content categories of the message, in order to define the way these elements could improve the capacity of the tutor to evaluate the progress of a discussion thread in a distance education forum.

Among others, the prediction for future research actions are long-term studies concerning the main issue: what reinforces the participation at fora and how this contributes to the educational process effectiveness by investigating side questions, such as how much it affects the person who starts the thread (tutor or student), how it starts, the period when the thread starts, how important the time of response in threads, is the groups' size etc. and their association with the elements concerning the students' profiles and their performance in course modules of HOU, intending to reach educational conclusions.

ACKNOWLEDGMENT

The authors would like to thank the anonymous referees for their careful review and useful suggestions.

REFERENCES

[1] A.W. Lazonder, P., Wilhelm, and S.A.W. Ootes, "Using sentence openers to foster student interaction in computer-mediated learning environments" Computers & Education, vol. 41, pp. 291–308, 2003.

[2] G. Stahl, "Building collaborative knowing: contributions to a social theory of CSCL" In What we know about CSCL in higher education (Strijbos, J.W., Kirschner, P., & Martens, R.L. Eds.), Kluwer, Amsterdam, 2004.

[3] C.N. Gunawardena, K. Carabajal, and C.A. Lowe, "Critical analysis of models and methods used to evaluate online learning networks" In American Educational Research Association Annual Meeting, American Educational Research Association, Seattle, 2001.

[4] F. Henri, "Computer conferencing and content analysis". In *Collaborative learning through computer conferencing* (KAYE AR Ed), pp. 117–136, The Najadan Papers, Springer-Verlag, London, 1992.

[5] D.R. Newman,, B. Webb, and C. Cochrane, "A content analysis method to measure critical thinking in face-toface and computer supported group learning" Interpersonal Computing and Technology, vol. 3, pp. 56–77, 1995.

[6] M. Bullen, "A case study of participation and critical thinking in a university-level course delivered by computer conferencing", University of British Columbia, Vancouver, Canada, 1997.

[7] P. Fahy, G. Crawford, and M. Ally, "Patterns of interaction in a computer conference transcript" International Review of Research in Open and Distance Learning, vol. 2, pp. 1-24, 2001.

[8] J.B. Pena-Shaff,, and C. Nicholls, "Analyzing student interactions and meaning construction in computer bulletin board discussions", Computers & Education,vol. 42, pp. 243–265, 2004.

[9] D. Lockhorst, W. Admiraal,, A., Pilot, and W. Veen, "Analysis of electronic communication using 5 different perspectives" Paper presented in a symposium conducted at the 30th Onderwijs Research Dagen (ORD), Kerkrade, The Netherlands, 2003.

[10] E. Zhu, "Meaning negotiation, knowledge construction, and mentoring in a distance learning course" In proceedings of selected research and development presentations at the 1996 national convention of the association for educational communications and technology. Available from ERIC documents: ED 397 849, Indeanapolis, 1996.

[11] S. Jarvela, and P. Hakkinen, "Web-based cases in teaching and learning: The quality of discussions and a stage of perspective taking in asynchronous communication" Interactive Learning Environments, vol. 10, pp. 1–22, 2002.

[12] A., Weinberger, and F. Fischer, "A framework to analyze argumentative knowledge construction in computer-supported collaborative learning" Computers & Education, vol. 46, pp. 71-95, 2006.

[13] A. Veerman, and E. Veldhuis-Diermanse, Collaborative learning through computer-mediated communication in academic education. In Proceedings of the Euro CSCL 2001, pp. 625–632, McLuhan institute, Maastricht, 2001.

[14] D. Cook, and J. Ralston, "Sharpening the focus: methodological issues in analysing online conferences" Technology. Pedagogy and Education, vol. 12, pp. 361–376, 2003.

[15] S. Schrire, "Knowledge-building in asynchronous discussion groups: Going beyond quantitative analysis", Computers & Education, vol. 46, pp. 49–70 , 2006.

[16] M. Kongovi, J. C. Guzman, and V. Dasigi, "Text Categorization: An Experiment Using Phrases", Lecture Notes In Computer Science, Vol. 2291, pp. 213-228, Springer-Verlag, Berlin, 2002.

[17] F. Sebastiani, "Machine Learning in Automated Text Categorization", *ACM Computing Surveys*, vol. 34, pp. 1-47, 2002.

[18] F. Sebastiani, "*Classification of Text, Automatic. Encyclopedia of Language & Linguistics*", 2nd Edition, Section: Applications of natural language processing, Vol. 14. Elsevier Science Publishers, 2006.

[19] R. Nock and F. Nielsen, "A Real Generalization of discrete AdaBoost", *Artificial Intelligence*, vol. 171, pp. 25-41, 2007.

[20] Y. Freund and R. E. Schapire, " A Decision-Theoretic generalization of on-line learning and an application to Boosting", *Journal of Computerand System Sciences,* vol. 55, pp. 119–139, 1997.

[21] S. Bloehdorn and A. Hotho, "Boosting for Text Classification with Semantic Features", Advances in Web Mining and Web Usage Analysis. Lecture Notes in Computer Science, Vol. 3932, pp. 149-146, Springer-Verlag. 2006.

[22] S. Kotsiantis, K. Patriarcheas, and M. Xenos, "A combinational incremental ensemble of classifiers as a technique for predicting students' performance in distance education" Knowledge Based Systems, vol. 23, pp. 529-535, 2010.

 ISBN: 978-1-61208-228-8

United Kingdom Local Authority Challenges in the use of Twitter and Other Social Media

Darren Mundy, Qasim Umer
University of Hull
Scarborough, GB
d.mundy@hull.ac.uk;q.umer@2008.hull.ac.uk

***Abstract*—Social media networks (such as Twitter) provide new opportunities for local government agencies to engage with citizens in their local communities. However, there are substantial challenges for such agencies to deal with in these social spaces. This paper provides a critical perspective on the challenges presented in the results of a focused research project on governmental communication through Twitter over a three month period from October to December 2011. The research presented in this paper, contributes to the growing number of research papers related to the effective use of social media platforms in governmental, organisational and other community spaces. Social media platforms are no silver bullet providing only another channel of communication with the customer. However, the identification of ways of using social media for transformative purposes in local government scenarios, are worth identifying and exploring.**

Keywords-social media networks; social network analysis; eGovernment; communication channels

I. INTRODUCTION

In March 2012, Facebook reported a worldwide audience of almost 836 million users [1]. In the same month, Twitter reported 140 million active users with 340 million tweets per day [2]. These statistics demonstrate the global presence of networks that have only been in existence for less than ten years. These global audiences may be dispersed, and not always balanced across local communities. With such large user bases, companies are exploring how they can best engage users on these platforms, and leverage a return on investment, from time committed to social media spaces.

Social media platforms such as those named above provide opportunities for individuals to collectively discuss, share, participate, produce, and respond to specific materials dependant on platform (e.g., videos on YouTube and short text messages on Twitter). Many of the social media platforms provide an opportunity for individuals to provide a picture of themselves (often through a profile), and opportunities to obtain and engage with information, often in real time.

UK government is managed through a mixture of national government departments and local government authorities. A primary requirement of the UK democratic system is a need to engage with, and listen to, constituents through a multitude of different access channels. Recently, with the austerity measures imposed across the whole of Europe, and other parts of the world, the UK has had to look again at the delivery of public services. These austerity measures have impacted on local government funding, meaning in places, cuts in front line services.

Organisations (including local government authorities) recognise the benefits that shifting customer enquiries from physical face to face customer contact centres, to online provision, can bring. These benefits include: reductions in the cost per transaction of customer enquiries; and where front end services are directly connected to back end services the potential for disintermediation. Social media platforms have been identified as platforms through, which to engage in providing front end service information, and in answering customer enquiries in cost effective processes.

Local councils in the UK, are present across a range of social media spaces with many individual councils engaging across multiple spaces. In addition to staff within local authorities engaging within social media spaces, many other publically funded organisations are also using these networks, as mechanisms to engage the local taxpaying public. This network of public sector engagement provides the citizen with direct access to queries about public sector services, and mechanisms for obtaining current localised information.

This paper builds on the research results presented in Mundy and Umer [3]. The previous paper presented the findings from a study focused on the engagement of ten local UK councils within the Twitter network. The paper outlined a range of quantitative statistics regarding the study and indicated a number of issues discovered within the textual analysis of the tweets. The aim of this paper is to explore the deeper challenges presented from the data, determined from detailed textual analysis of the tweets collected for the study. Whilst the study focused on Twitter many of the challenges presented within this paper, can be easily applied to other social media networks. The paper contributes to a growing number of research papers related to the effective use of social media platforms in governmental, organisational, and other community spaces.

This introduction has provided an overview of the importance of social media networks, and the potential for social media networks to impact on the government: citizen relationship. Section II gives an overview of literature in the area of social media communication, particularly highlighting the effective use of social media in government spaces. This is followed by a detailed review of the methodology used to analyse local government conversations in the 'Twitterverse'. These conversations are then used to identify clear challenges in the use of Twitter, and other social media platforms to engage the citizen.

 ISBN: 978-1-61208-228-8

Finally, a series of recommendations linked to the challenges are provided and the paper closes with a conclusion and the identification of further work.

II. SOCIAL MEDIA AND UK GOVERNMENT

The growth of social media is impacting the ways in which communities work, communicate, and socialize. According to Landsbergen [4] social media platforms can help to fulfill the needs of rapid communication, to engage individuals with multimedia artefacts and problematically to blur what is private and what is public. Similar to the demands imposed by users of the short messaging service (SMS) on mobile phones, social media networks demand prompt response, knowledge of and continued engagement with the platform. Twitter, Facebook and YouTube are examples of three multilingual social networking websites launched since 2004 that respectfully provide forums for social networking, micro blogging, and multimedia content sharing (text, video and photos). They give users the opportunity of being real time in a virtual world [5], and enable users to create their own accounts, content channels and interest group sites. Different governments, channels and groups also have their own individual pages and accounts on social media networking sites.

This paper focuses on the use of and challenges in the use of social media, in particular, Twitter by local government in the UK. Research into the use of social media networking sites by national governments, and organisations across the world continues to grow, as we continue to look for how transformative communication can be provided through such channels. Recently, Stephen Goldsmith used the term 'government by network' [6] to describe how online communication channels were being used at different levels of government, to engender a greater sense of participative relationship with citizens. In addition, there is evidence that social media networks when employed effectively have the power to create change in relation to political discourse, encouraging individuals to re-engage with democratic systems [7] and create a greater sense of the citizen voice [8].

Research in the area of government use of social media networks can be broken down into two primary areas of interest:

- Analysis at national level of how parties and political leaders utilise these communication channels to engage citizens with national and international political issues.
- Engagement with how local government organisations are using such platforms to engage citizens with local services and issues.

Research related to local government usage of Twitter is not as well established as national level research, as the national issues often generate more substantial interest. In this national space, the most interesting study to date has focused on how Barack Obama (current president of the USA), utilised the web and social media networks in his historic election victory in 2008 [7]. Other researchers have focused on an analysis of the use of social media networks in relation to encouraging political dialogue [9], the use of social media for political public relations [10] and analysis of social media channels as political communication channels [11]. Researchers have found from a national perspective that broadcast information over these channels is broadly favoured, and that individual channels are not in themselves 'game changers', but merely an additional communications channels for contact with citizen groups [9]. From a national perspective, it is useful to note Tumasjan et al. [11], which suggests that these network channels can be used to provide a prediction on the results of national elections. This in itself is not unique (given perhaps we could also use analysis of party prevalence in other forms of media or surveys), it does suggest that political discourse at the national level is frequent, and the size of it is nationally relevant on social media channels.

Whilst national issues are of interest this paper focuses primarily on investigating local government discourse. One of the key issues for local government is engendering citizen interaction in positive, and progressive ways. Social media platforms can offer opportunities for individuals within local communities to provide their view on a local issue, report a broken street light, or to simply interact with a local councillor. Researchers such as [6][12][13] have indicated that these channels can enable a transformation in the way in which local government, and citizens communicate leading to transformative relationships developing. In particular, Danis et al. [13] argues that these networks can enable "an atmosphere of co-operation" as citizens work with local government in developing better physical communities.

A barrier to the co-operative atmosphere highlighted in the above paragraph, arises in the ways in which local government utilise social media spaces. Rooksby and Sommerville [14] suggest that the ways in which many local government organizations, manage their information technology infrastructure, and the ways in which local government service structures are established, can limit the ways that such organisations are able to utilise social media systems. They state within these circumstances the use of social media channels *"can only practically be used to broadcast announcements rather than to interact with people"*. As a practical example one could posit the traditional ways in which customer service interaction, and public relations are managed in local government organizations, can often lead to uncertainty with regards to the translation of such services on to social media platforms.

Noting the growth in usage, by local government of social media, recently research has been published to try to provide such organisations with a framework of good practice [15]. However, this area continues to evolve and further work is needed to understand where local government is now with its engagement in social media spaces, what the challenges are with regards to this engagement, and how such engagement can continue to evolve. This paper represents work towards establishing a response to the above, highlighting in particular, the challenges to local authorities effective use of social media platforms such as Twitter, Facebook and YouTube.

 ISBN: 978-1-61208-228-8

III. METHODOLOGY

Over the period October – December 2011, tweets related to ten local councils in the UK, were analysed. The councils selected for analysis were drawn from a group of UK councils with active accounts within Twitter. Therefore, the analysis focused only on those councils who were experimenting with this social media network, and were already engaging in some way, with their local populace in the 'Twitterverse'.

Tweets were collected from within defined council spaces and from outside of defined council spaces. This effectively translates through to collection of tweets that each council had made, collection of citizen tweets to the council, and tweets made which made reference to the council or the local community. It should be noted that no collection was possible for those communications happening within Twitter through private communication channels (e.g. through Twitter's direct messaging system).

The aggregated data collected provided quantitative metrics covering aspects such as:

- The number of tweets;
- The number of followers;
- The number of re-tweets etc…

In addition, the text within each tweet was analysed in relation to:

- Categorising the focus of the content;
- The direction of the communication (e.g., whether it was just broadcast or a response to another Twitter user);
- An indication of the feel of the tweet (e.g., whether it expressed positive, negative or neutral emotions);
- A detailed content analysis of the messages contained in the tweet content.

In relation to the detailed analysis of the messages in the tweet content, each tweet was read, and broader issues were identified within the tweet content. This paper provides a critical analysis of these broader issues, and identifies the challenges presented by these.

There were a number of limitations presented in the research data gathered within this project. The main issues were linked to the choice of councils, the time of year studied, and the maturity of various councils' use of Twitter as a communication channel. However, these limitations have limited impact on the broader challenges presented through this analysis work.

In addition to this direct identification of challenges from analysis of the research data, a reverse approach has also been used. Conversation with individuals responsible for social media management in local authority spaces has helped to inform and identify other challenges present within the research dataset.

In total, 1,565 tweets, were analysed from within council defined spaces in Twitter. In addition to this, a substantial number of tweets were also reviewed in spaces outside of council control. Content from the tweets analyzed can be found throughout Section IV in providing evidence, to support the challenges identified.

Mundy and Umer [3] provides initial analysis of the study detailed above. General information about the study is presented including the number of tweets analysed (n= 1565), no of followers for each council, and total number of re-tweets made in the sample set (n=191). Detailed quantitative analysis is also provided regarding categorisation of the content of each tweet and the directionality of the communication.

It was found that the most popular reason for councils to make announcements through Twitter is to publicise entertainment, sport or leisure activities occurring in the local area. In addition, the focus of re-tweets made by local councils centred on those services that individual councils wished to highlight most in their local area. For example, Newcastle upon Tyne City Council highlights local library services, Salford City Council highlights local health related services, and the City of Lincoln council highlights local business services. The ways in which many councils engaged with responsive texts were fairly limited, or in some cases non-existent. The most popular categories for responsive tweets from UK Councils were centred on transportation, entertainment, waste and housing services.

Finally, within Mundy and Umer [3] analysis is provided of a range of conversations which occurred within the council Twitter spaces. These conversations point to issues regarding the type of conversation which occurs through Twitter, the presence of offensive content, Twitter best practice, and individual privacy issues. This follow on paper focuses on a deeper critical analysis of the challenges presented within qualitative Twitter content analysed in the earlier study.

IV. LOCAL AUTHORITY CHALLENGES

The next few sections will take an in-depth look at challenges identified in the analysis of the tweets, contained within the study.

A. Transforming Broadcast Communication

As noted by Rooksby and Sommerville [14], there are significant challenges in transforming different elements of local government services, away from broadcast only models of communication. Social media communication platforms provide direct public channels of communication between citizens and local government. This can lead to citizens directly criticising decisions made in local communities, and the transparency of local government [3]. It can also lead to citizens raising issues about local service provision or issues related to local areas. The problem is how to construct these social media channels such that appropriate individuals receive the messages related to their particular services, and how to manage this communication within these channels. In other circumstances, often the conversation is less public and more easily controlled – this leads to local government organisations (similar to other commercial organisations) approaching two way conversation in social media spaces, with some trepidation. However, where local government is engaging directly in this form of practice, there is ample evidence that citizens are appreciating this service, with some commenting that social media platforms are the most

 ISBN: 978-1-61208-228-8

effective ways in which they have communicated with government services.

B. Frontend to Backend Integration

Local government must take clear steps to manage the social media channels they are using from the perspective of understanding information flow. Tweets show evidence that for some councils citizens see their twitter spaces as the most effective way to gain answers to questions posed. However, there were a number of cases of citizens not obtaining the information they required. One solution may see customer service staff taking responsibility for the engagement in answering citizen questions, within social spaces, and other specific staff (e.g., marketing) taking responsibility for highlighting council achievements. In addition, customer requests coming in through social media spaces, must be tracked to help to facilitate knowledge of return on investment, as highlighted in Section III.C.

C. Leveraging Return on Investment

With government austerity measures in force, and a reduction in funding to local government in the UK, local councils are busy reducing inefficiencies. Therefore, for social media channels to be embraced they must make clear demonstrable impacts on local citizen engagement. One way in which this can occur, is by moving transactions from more expensive channels (e.g., face to face services) to online information provision. This is where it is important to ensure that the social network spaces are joined up to local government services; otherwise, the cost of a transaction and speed of reply may be similar to other channels particularly if answers are not clear or incorrect. The most effective councils using these platforms make use of the private messaging spaces within social media platforms to provide targeted complete answers to citizen questions.

D. Engaging the Citizen

Engaging the citizen will be helped by a transformation of broadcast communication, but this is not the only challenge to providing an engaging service. Social media platforms are often at their best when the channel providers are utilising the channel effectively by posting new interesting content. Often, individuals will post content that is delivered across multiple social media channels this can cause problems when users are signed up for multiple services, each service needs to be distinctive, and different to engender different user communities. Whilst platforms can be transformative in providing a greater relationship between citizen, and government, they can also be transformative in a negative way, if the use of the channel does not meet with citizen expectations; for example, if the channel is used to broadcast less meaningful information.

E. Privacy Challenges

Social media platforms are not private and often citizens within these spaces are not completely anonymous. This presents a clear challenge in ensuring details of customer cases, are kept out of the public domain. In one particular poor example of this, a council engaged across multiple tweets in detailing reasons for action, in a specific individual's case. This could be deemed to infringe on data protection when specific details are discussed.

The counter concern is where customers are willing to discuss their individual personal circumstances across public spaces. Councils must learn when to move conversation into more private spaces, and to find ways of educating local citizens about problems related to providing too much personal information within public social media spaces.

F. Engaging with the Conventions

Social media spaces are not the same, all have particular nuances, and the community within different spaces behaves often in particular ways. Understanding the nuances of the platform will lead to greater success in platform use. For example, Twitter encourages users to follow other users and build networks through interconnections of individuals, the community also has established conventions like #FF, which is short for #FollowFriday (Twitter users use this at the end of a week to suggest other individuals to follow). Recognising, understanding and using the structures available can help councils to fit into the community and be seen as an interesting component part.

G. Small Numbers

At present, council social media spaces have relatively small numbers, in terms of percentages of citizens within local communities engaging with the service offered by and through them. However, as councils gain a better understanding of how to use and market their presence on social media networks effectively then these numbers should grow. It may well be for certain community groups engagement through social media spaces is the only way to engage these groups in conversations about local council spaces, and services.

More mature use of social media within council spaces is looking directly at how networks are forming in social media spaces, around the council. Understanding the community in terms of the different types of user present, and engaging in the space, is extremely important. Equally, using the community as a network to advertise the strength of the service should be one thing councils look to do.

H. Endorsement or Community Building

It is evident from the research that the practices that councils are using in building communities within social media spaces differ. Decisions such as whether the council should 'follow' other users or 'join' other groups are not always that simple. However, social media networks are often formed from practices of engaging in community building. For example, presence as a friend may highlight your presence to other users who are interested in you. Similarly engaging in the process of #FF, or re-tweeting content, also provides a sense of social media community engagement, and may influence others in passing on your content. Another example is the highlighting of other videos for consumption within your YouTube channel. However, these community ideals come at a risk, the risk that the content you highlight, are associated with, or pass on is

 ISBN: 978-1-61208-228-8

inappropriate or legally problematic. There is a question over whether associating local councils with other information or users provides a form of endorsement of them or their services, and therefore, whether such endorsement is appropriate given the linkage to public services.

I. Answering the Difficult Questions

Social media spaces provide direct lines of communication to the general public, which are open 24/7 and easily communicable. These spaces are used by individuals for a variety of purposes from communication about particular things happening within local communities, through to information discovery about other places before travel. This creates complications regarding the types of issues raised within these spaces. In many ways, they are the most transparent places for citizens to find out more about what is happening within local communities direct from other citizens. This obviously means that some of the positives will be highlighted, but it also means that many of the issues often kept outside of general knowledge may also be highlighted, and associated with the spaces. For example, conversations about local drug use, anti-social behavior, areas of violence, and other criminal activity.

J. What is the difference?

Many of the above points raised as challenges for local authorities in these spaces can also be linked directly to challenges for large organisations. However, the interesting question is whether there is anything that distinguishes the challenge for e-government. In this area, there are thought to be two major differences.

The first is the increased transparency and profile of the service, money to fund engagement of councils within social media spaces comes directly from the UK public purse. The service needs to have a clear demonstrable impact for citizens within the local community to deem this as successful.The UK press are quick to highlight customer service faillings in online spaces; take, for example, Tameside Council's experimentation with a virtual customer helpdesk in Second Life described as "absolutely barmy" [16]. Although the value has to be demonstrated in large private organisations, they are not forced to be as transparent; for example, with detail regarding cost of service.

It is interesting to see citizens commenting on the state of local government transparency through social media spaces, e.g., "#organisation is in breach of international law, but wants the £4.7bn contract for N London waste. Tell #Camden council 'No' #Palestine #Gaza@" and "People getting turned away... doesn't feel like an 'open council' tonight. #newcastle" These suggest that UK citizen's are willing to use the service directly to let councils know when they seem to be doing things thought not to be in the public interest.

The second percieved significant difference is that councils could be deemed to have a requirement to engage the citizen, whether this is in citizen democracy, or in conversation regarding quality of services. Other organisations may choose to engage customers but local councils need to engage citizens. Therefore, understanding where particular customer groups are conversing, and using up to date mechanisms for engaging citizen groups should be a clear part of the strategic engagement for all local councils.

V. CONCLUSION AND FUTURE WORK

The challenges outlined above demonstrate that councils need to think carefully about the ways that they use social media channels to engage the citizen. Arising from the analysis, recommendations can be formed as to what councils should consider in having a presence, within these spaces. The recommendations themselves are not new but it is evident from the research that councils are not completely engaging in understanding, and applying, the recommendations in practice.

Understand the channel – Each channel has particular nuances, those councils who understand how these work seem from the data analysis to be perceived more preferably within the social media community. In addition, those demonstrating an understanding of the channel seem to obtain the greater amount of participation and the greater amount of two way communication.

Engage the citizen – From the research there was evidence that some councils are still not engaging in any form of two way dialogue, within social media spaces. Those councils who were engaging in a conversational manner with constituents, seem to have been generating a positive response to their social media engagement.

Develop policy – It was evident within the research that few councils had developed a charter for their and their citizens engagement within social media spaces. Developing clarity over what is and what is not acceptable in the spaces, may form a barrier to conversation but would provide a clearer sense of the general rules of engagement within the space. One argument may be that the spaces themselves often have guidance over what should and should not occur within the social media channel, however, in some cases, this may not be restrictive enough.

Advertise the channel – The best form of advertising in social media spaces, is the citizen. However, many organisations are finding innovative ways to highlight their involvement within the spaces. For example, Marseille FC offered their fans the opportunity to design a shirt if they gained a set number of Facebook followers. The best advert for any social media channel is the value added to customer experiences. If the channel is perceived as useful then it is likely that individual users will pass on that information, whilst if the channel provides in the main useless information, then it is likely to be ignored.

Integration – As outlined as a channel above, integration is extremely important for local councils. Managing the ways in which information flows into (in the form of physical posts), and out of the social media space is crucial (in the form of citizen driven requests or data). Social media use should not be in the hands of a singular person tagged with the responsibility of a social media producer, or equivalent. However, the ways in which councils, are represented within the space, is extremely important. There is a clear balancing act between bring overly prescriptive and transparent (e.g.,

 ISBN: 978-1-61208-228-8

tagging each post with a service representatives name) through to allowing for total freedom of employee engagement within the space. In addition, there are significant questions, which arise linked to whether engagement within social media spaces should be part of the role of senior figures such as chief executives of local councils.

Withdrawal – It is evident over the past ten years that developed spaces will fall in, and out of popularity. Early providers of social media platforms are finding reduced network engagement over time whilst newer platforms have come in, and filled their spaces. Therefore, councils need to manage their portfolio understanding how decisions are made to engage within particular spaces and when decisions should be made to withdraw from engagement within the space. This is a particularly complicated problem if there are constituent groups who maintain loyalty to particular social media spaces. In this space, councils should also consider audit and control mechanisms. If particular networks were to be forced to close it would be problematic if materials stored within those spaces were not stored in one form, or another. In addition, ensuring that any information that would generally be stored for legal purposes in the physical space, can be retrieved when engaging in virtual spaces.

The analysis, which led to the indication of challenges for local councils in engagement in social media spaces, did in itself, demonstrate a lack of lack of true engagement at present in these spaces. A particular problem for councils is moving away from the broadcast model given the often fragmented nature of services, and the traditional customer service mechanisms used. The analysis highlighted that councils need to do more to understand the types of data customers want to see within these spaces, and a need to understand more appropriate strategies for managing public:private communication through the different channels. These findings should be of use and relevant to national and international managers of e-Government web services, government legal teams and senior managers in eGovernment.

REFERENCES

[1] Internet World Stats, "Facebook Growth Stats", http://www.internetworldstats.com/facebook.htm 03.08.2012.

[2] Twitter Blog, "Twitter turns six", http://blog.twitter.com/2012/03/twitter-turns-six.html 03.08.2012.

[3] D. Mundy and Q. Umer, "An Analysis of UK Council use of the Social Network – Twitter," Proceedings of the 12th European Conference on E-Government, Barcelona, 2012 pp. 502-511.

[4] D. Landsbergen, "Government as Part of the Revolution: Using Social Media to Achieve Public Goals," Electronic Journal of e-Government, Volume 8 Issue 2, 2010, pp135-147.

[5] J. J. Zhao and S. Y. Zhao, "The Impact Of E-Government Use Of Social Networking Media On Government Transparency And Public Participation," Télescope, Vol. 18, 1-2, 2012, pp. 44-61.

[6] R. Cole, "Social media: what does it mean for public managers?," PM Magazine, Volume 91, Number 9, 2009.

[7] J. Hendricks and R. Denton, Communicator in Chief, Lexington Books, 2010, ISBN-13: 978-0739141052.

[8] J. Skinner, "Social Media and Revolution: The Arab Spring and the Occupy Movement as Seen through Three Information Studies Paradigms," Sprouts: Working Papers on Information Systems, 11(169), 2012.

[9] G. Gulati and C. Williams, "Communicating with Constituents in 140 Characters or Less: Twitter and the Diffusion of Technology Innovation in the United States Congress," Presented at the Midwest Political Science Association, Chicago, Illinois.

[10] R. Waters and J. Williams, "Squawking, tweeting, cooing, and hooting: analyzing the communication patterns of government agencies on Twitter," Journal of Public Affairs, Volume 11, Issue 4, 353–363, 2011.

[11] A. Tumasjan, T. Sprenger, P. Sandner, and I. Welpe, "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment," Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, D.C., 2010, pp. 178-185.

[12] R. Romsdahl, "Political Deliberation and E-Participation in Policy-Making," CLCWeb: Comparative Literature and Culture, Purdue University Press 7(2), 1-1, 2005, pp. 1-11.

[13] C. Danis, M. Bailey, J. Christensen, J.Ellis, T. Erickson, R. Farrell, and W. Kellogg, "Mobile Applications for the Next Billions: A Social Computing Application and a Perspective on Sustainability," 2nd Workshop on Innovative Mobile Technology and Services for Developing Countries (IMTS-DC 09), 30-31 July, 2009, Kampala, Uganda.

[14] J. Rooksby and I. Sommerville, "The Management and use of Social Networks in a Government Department," Journal of Computer Supported Cooperative Work (CSCW), October 2012, Volume 21, Issue 4-5, pp 397-415.

[15] N. Williams, "Template Twitter strategy for Government Departments," http://neilojwilliams.net/missioncreep/wp-content/uploads/2009/07/17313280-Template-Twitter-Strategy-for-Government-Departments.pdf 03.08.2012.

[16] The Telegraph, "Council scraps £36,000 virtual town hall in second life", 21 July 2010, http://www.telegraph.co.uk/technology/internet/7901312/Council-scraps-36000-virtual-town-hall-in-Second-Life.html 03.08.2012.

 ISBN: 978-1-61208-228-8

Technical Aspects of Sustainable Digital Archives

Cheng-Hung Li[1], Wei-Chun Chung[1], Shiang-An Wang[1], Chia-Hao Lee[1]
Chi-Wen Fann[1], Lee-Hom Lin[1], Jan-Ming Ho[1,2]
[1]Research Center for Information Technology Innovation
[2]Institute of Information Science
Academia Sinica
Taipei, Taiwan
{chli, wcchung, sawang, chahao, fann, monica, hoho}@iis.sinica.edu.tw

***Abstract*—In Taiwan, the National Digital Archives Program was initiated in 2002. It was subsequently integrated with the National Science and Technology Program for E-Learning into the Taiwan e-Learning and Digital Archives Program (TELDAP) on January 1, 2008. The program has created hundreds of professional digital archives, websites, and databases to store and manage these valuable national cultural and ecological artifacts. These websites and databases are open to the public. TELDAP will celebrate its 10-year anniversary and recess at the end of 2012. In this paper, we will present efforts made by the Digital Archive Architecture Laboratory, Academia Sinica, for preserving and maintaining accessibility of contents through previously developed systems. We will primarily focus on the reliability of the archive system and the promotion of its diverse contents. Our long-term preservation procedure includes both data and software. In the area of preservation, we recommend digitization guidelines for data formats so that these valuable artifacts can be digitized at a level that is equivalent to state-of-the-art display quality. We will also provide format conversion services to convert digital objects to the latest common open format. The digital archive developed and proposed 9 standard operating procedures for each stage in the development of the digital archive system. In practice of software maintainability, we will preserve the design documents at each stage of development. Currently, we are developing a management framework on the basis of distributed virtualization technology to host digital archive systems so that management cost can be minimized while systems can remain operational.**

Keywords-digital archive; digital preservation;long-term preservation.

I. INTRODUCTION

Preservation of human knowledge for long-term accessibility is one of the missions of an institution of cultural and ecological heritage. For thousands of years, paper has been the preferred medium for storage of text and images. Currently, due to the rapid development of information technology, numerous objects (e.g., books, cultural relics, paintings, calligraphy, etc.) have been archived in digital form for their dissemination, backup, and reuse [1][2]. However, whether institutions decide to use either analog or digital approaches to archiving, they will need to employ comprehensive preservation strategies for sustainability so that archived objects will remain readable [3][4].

The TELDAP [5] is a joint project between scholars in the fields of humanities and technology. It has built a platform known as the portal of TELDAP [6] to preserve the artifacts of Taiwan's rich cultural heritage in digital form. This program functions through the collaborative efforts of 19 organizations and government agencies. It has archived over 5 million digital objects, along with metadata annotated by domain experts. In addition, it has created hundreds of professional digital archives, websites, and databases to store and manage these valuable national cultural and ecological artifacts. We must address three types of problems that affect the preservation of these systems: physical deterioration, technology obsolescence, and improper management. The first two problems are similar to problems that occur during traditional digital preservation. Both processes involve difficulties related to the preservation of format materials. In addition, inadequate management of existing systems can cause gradual system loss. In the past, when archiving institutions outsourced their systems, they did not insist on comprehensive analysis and planning of system requirements. Most of these systems lacked architectural flexibility. Therefore, expansion and data integration of these systems were difficult. Some systems had to be closed due to a lack of standard maintenance and development processes. We hope to improve the reliability, overall representation, scalability, quick response, and user-friendly, value-added environment for these valuable archives through the development of a complete suite of standard operating procedures (SOPs) that can be used to fully integrate and preserve these websites, systems, and databases.

In this paper, we propose a plan that details preservation strategies for digital archives. Through the standard process of formulation, an archive system can be developed through standardization. Thus, the system can be preserved completely through short-term system migration or long-term preservation.

This paper is organized in the following manner. In Section II, we review literature related to our plan. We also review relevant strategies for archive system preservation. In Section III, we provide detailed explanations of architecture of the sustainable digital archives we have developed. Finally, in Section IV, we present our conclusions.

 ISBN: 978-1-61208-228-8

II. RELATED WORK

In this section, we introduce several issues and techniques that relate to the preservation of digital information. Digital Preservation involves a set of processes, activities, and management of digital information that is conducted over time to ensure the long-term retention and accessibility of archived information. Hence, it is not just the product of a program. It is an ongoing process. Functions within this process include management of object names and locations, updates of the storage media, and documentation of content and tracking of hardware and software changes to ensure availability and comprehensibility of objects [7]. The Research Libraries Group [8] defines Digital Preservation as a series of activities conducted to ensure that digital data can be maintained and queried continuously. The American Library Association (ALA) briefly defined Digital Preservation as a series of policies, strategies, and actions employed to ensure access to the digital contents over time, despite the challenges presented by media failure and technological change. Although there has been progress in the field of Digital Preservation, many tasks remain undone: Professionals must decide which types of file formats should be preserved and agree on the level of preservation needed. Further, in order to promote responsible stewardship of digital information, professionals must also agree to comply with standards. Presently, data can be preserved for as little as five years or for as long as more than ten years [10]. In fact, some scholars believe that the preservation period should last more than a lifetime.

Therefore, it is evident that the accessibility of digital information is extremely vulnerable in today's rapidly evolving technological environment. A majority of the selected research stated that the field of digital preservation faces three major obstacles: physical deterioration, technology obsolescence, and the rapid growth of information. The problems inherent in the preservation of archival objects in both digital and traditional print form are rather distinct. Therefore, preservation strategies must be studied and evaluated; further planning is necessary [11]. The National Archives of UK suggested a number of criteria for selection of file formats for long-term preservation: ubiquity, support, disclosure, documentation quality, stability, ease of identification and validation, intellectual property rights, metadata support, complexity interoperability, viability, and re-usability [12]. In 1995, in order to address the issue of Long-term Digital Preservation, the International Standard Organization (ISO) established The Consultative Committee for Space Data Systems (CCSDS). In 2003, this committee formulated the Open Archival Information System (OAIS, ISO14721:2003) to provide a complete reference model of digital preservation architecture [13].

Knowledge storage institutions face a common dilemma: How can digital information sustainability be preserved? In the fight against the loss of digital information, several technical approaches have been proposed: refreshing; migration; emulation; standardization; system (technology) preservation; encapsulation (data description); replication (redundancy); and conversion to paper or analog media. Several researchers have examined migration, emulation and system preservation as strategies for digital preservation, [14][15][16][17]. Each of the aforementioned traditional approaches or strategies suggest that preservation of information can be achieved by the use of technology.

In addition, the aim of the research of the Software Sustainability Institute was to raise awareness and build capacity throughout the Further and Higher Education (FE/HE) sector through engagement in preservation issues as part of the process of software development [18]. They provided a breakdown of the different approaches used in the achievement of software sustainability: technical preservation (techno-centric), emulation (data-centric), migration (functionality-centric), cultivation (process-centric) hibernation (knowledge-centric), deprecation, and procrastination. The success of each of these approaches depends on many factors: the importance of the software, the maturity of the software, the size of its community, and the availability of resources needed to achieve sustainability. The suggested framework can help groups of developers to understand and gauge the benefits of ensuring that preservation measures are built into software development processes and the benefits of active preservation of legacy software.

III. OUR ARCHITECTURE

The purpose of TELDAP is to build a flexible architecture that emphasizes interoperability, continued maintenance, and continuous development. Most of TELDAP's digital archives systems are software-based projects that include four elements: hardware, system software, application, and data. Based on the properties of these elements, projects can be divided into different sustainable levels for each element (e.g., service sustainability, software sustainability, and data sustainability). In Figure 1, the hardware and system software are categorized under service sustainability and the application and data are categorized under software sustainability and data sustainability, respectively.

In order to ensure digital archives sustainability, we composed a set of rules to manage the processes of a given software-based project. We formulated complete Standard Operating Procedures (SOPs) for digital archives sustainability for different levels and tasks: Digitization (DP); Development of Software Project (DSPP); Metadata Interoperability (MIP); Intellectual Property Inventory (IPIP); User Account and Single Sign-on (UASSP); Website Identification (WIP); Website Traffic Observation (WTOP); Catalog Index and Data Exchange (CIDEP); and Long-term Service Preservation (LSPP). As we can see in Table 1, each task among the different levels of digital archives sustainability needs different SOPs to maintain its integrity and interoperability. The service sustainability level includes hardware and system software. Therefore, the procedures of LSPP, WTOP, and UASSP must be considered when this project needs to host, validate, or migrate. The main procedure in software sustainability level is DSPP, which helps the content provider control the development flow of software. For data sustainability, the MIP and CIDEP are

 ISBN: 978-1-61208-228-8

major procedures used in validation and migration tasks. Table 2 shows the combination of the framework of the Software Sustainability Institute and OAIS with our proposed architecture. In this paper, we will discuss DSPP, CIDEP, and LSPP in software sustainability, service sustainability, and data sustainability, respectively.

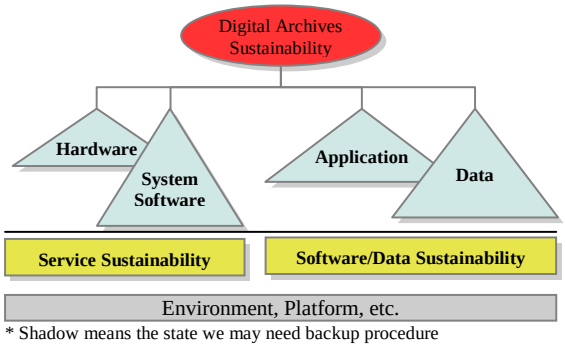

Figure 1. The Architecture of Sustainable Digital Archives

TABLE I. THE SOPS WITH THE DIGITAL ARCHIVES SUSTAINABILITY

Levels / Task	Service Sustainability	Software Sustainability	Data Sustainability
Hosting	LSPP WTOP UASSP	Documents (Application)	Documents (Data format) IPIP, WIP
Validation		DSPP	MIP, CIDEP
Migration			

TABLE II. THE SOPS WITH THE FRAMEWORK OF SSI AND OAIS

Subjects / Task	Service Sustainability	Software Sustainability	Data Sustainability
Hosting	Emulation	Hibernation	
Validation		N/A	OAIS
Migration		Migration	

A. *Development of Software Project Procedure (DSPP)*

In this section, we aim at discussing the improvement of the quality and performance of the software development process and ensure flexibility, scalability, and completeness for the digital archive system. As shown in Figure 2, we applied the Verification and Validation in process area of CMMI—level 3 and ISO/IEC 15504 in order to integrate the development flow of software project with the participant and related output development documents. This procedure defined the development flow and documents to help the content providers (Digital Archives Systems) control the processes in each step and confirm the results when completed by the plan and development groups.

B. *Catalog Index and Data Exchange (CIDEP)*

TELDAP continuously archived over one million digital objects and approximately 100 websites with databases. In the process, TELDAP digitalized a large amount of data that contained information on a variety of cultural heritage, historical files, and archaeological artifacts. These are distributed as heterologous website systems. The task of integration, preservation, and popularization of these cultural objects from different websites that were constructed by different institutes is both important and enormous. As shown in Figure 3, we implemented an integrated platform for the cross-directories' knowledge retrieval platform on the basis of an intelligent crawler that used the behavior analysis of the digital archive users.

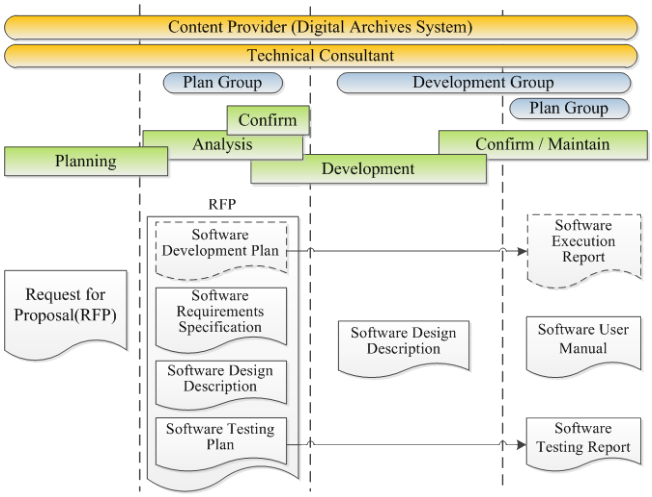

Figure 2. The development flow of software project

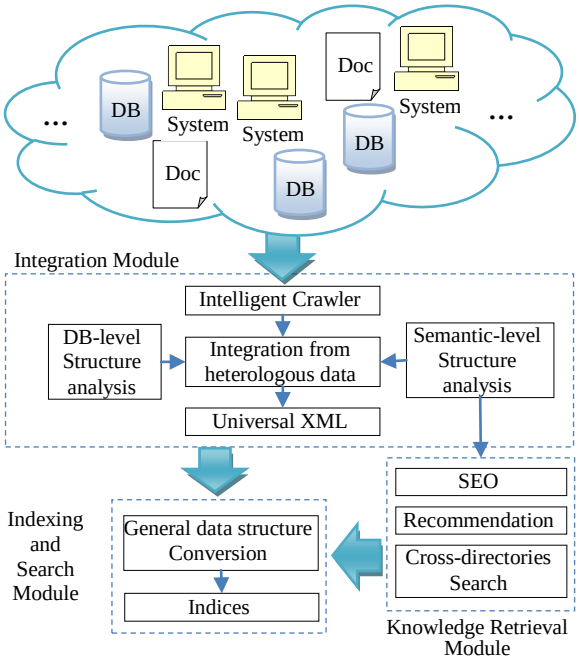

Figure 3. The architecture of the catalog index and data exchange

This framework used an intelligent crawler and the integration of heterologous data, indexing and search, and knowledge retrieval. In addition, it presented a user interface that offers search objects to the general public and provides opportunities for the public to study the information in these archives. Based on the achievements of this framework, individuals can easily enjoy the treasures of 5000 years of Chinese civilization. They can also learn about Taiwan's ecology and history.

 ISBN: 978-1-61208-228-8

C. Long-term Service Preservation (LSPP)

In system preservation, we must consider website migration, duplication, and emulation. Migration is used to ensure usability between generations of media. Migration also insures that a content object can be sustained in new generation media. Duplication is usually employed for the fault-tolerance of contents. Creation of an off-site backup can improve error recovery and ensure the availability of content. We believe that duplication is a basic element of preservation. Through emulation, a system can be maintained in a stable state. There are two types of emulation: static and dynamic. For a website that contains static pages, preservation by *snapshot* is recommended. When a crawler is employed, the contents of the site can be retrieved and saved by using a particular format.

For the dynamic-type system (e.g., page content within a database), we recommend the use of emulation or virtualization. An emulator is used to simulate hardware devices on the software platform. It causes the guest Operating System (OS) to run as a Personal Computer (PC). Virtualization, a state-of-the-art computer science strategy, is used to allow a guest OS to access (and/or share) hardware with a monitor (a hypervisor) on a host's OS. The choice of strategy involves trade-offs between performances, flexibilities, scalabilities, security, etc. Combinations of migration, duplication, and emulation can be used to fulfill the many requirements of digital preservation. As Table 3 indicates, these combinations can simultaneously meet the needs of a variety of websites and their availabilities.

TABLE III. STRATEGIES OF PRESERVATION

Method			Description
Duplication	Emulation	Migration	
V			The foundation of digital preservation
V	V		Snapshot: For static-typed website Virtualization: For dynamic-typed website
V	V	V	Ensure contents can be read in the future

IV. CONCLUSION AND FUTURE WORK

In this paper, we proposed strategies for the improvement of the reliability, overall representation, scalability, quick response, and friendly value-added environment for valuable archival objects. We hope to assist institutions involved in digital archiving to formulate a standard process for system analysis, planning, development, testing, and acceptance. In addition, we aim to provide technical consultation to these institutions to help them provide specifications and advice to outsourced vendors on the development of a standard digital archive system. We also provide guidance on the best way to draft a document to outline standard digital archiving procedures. Overall, we have provided a number of strategies that can help digital archiving institutions develop, manage, and preserve digital archive system sustainability

However, there are several areas for improvement. Until now, digital archiving institutions have faced many limitations in the areas of operation and innovative techniques. We need to develop new preservation strategies to address these new technologies. We must also provide ways for institutions to improve their operations so that they synchronize with standard processes. In the future, we hope to provide more developed plans for the improvement of the interoperability and integration of the wide range of available archiving systems. In doing so, we hope to preserve the integrity of current information for use in the future.

ACKNOWLEDGMENT

This research was supported in part by Taiwan E-learning and Digital Archives Programs (TELDAP) sponsored by the National Science Council of Taiwan under NSC Grants: NSC101-2631-H-001-005, NSC101-2631-H-001-006 and Academia Sinica.

REFERENCES

[1] M. Hedstrom and S. Montgomery, Digital Preservation Needs and Requirement in RLG Member Institutions, Report, A Study Commissioned by the Research Libraries Group, Dec. 1998.

[2] F. Moore, "Long term data preservation," Computer Technology Review, Third Quarter 1999, pp. 32-33.

[3] D. A. Kranch, "Preserving electronic documents," Proceedings of the Third ACM Conference on Digital Libraries, May 1998, pp. 295-296.

[4] E. Cooly, and N. Chip, "Integrating solutions: examining the collection management process using OCLC's firstsearch electronic collections online," Library Acquisitions: Practice & Theory, Vol. 22, No. 1, 1998, pp. 97-102.

[5] Taiwan e-Learning and Digital Archives Program, *http://teldap.tw*.

[6] TELDAP Resources Portal, *http://digitalarchives.tw*. 08.15.2012

[7] DigitalPreservationEurope, *http://www.digitalpreservationeurope.eu*. 08.15.2012

[8] RLG/OCLC, "Trusted digital repositories: attributes and responsibilities: an RLG-OCLC report," RLG/OCLC Working Group on Digital Archive Attributes, May. 2002, [Online]. Available: http://www.rlg.org/en/pdfs/repositories.pdf. 08.15.2012

[9] American Library Association, Definitions of Digital Preservation. Definitions of Digital Preservation, 2007, [Online]. Available: http://www.ala.org/alcts/resources/preserv/defdigpres0408. 08.15.2012

[10] I. Verheul, Networking for Digital Preservation. Current Practice in 15 National Libraries, IFLA Publication Series, [Online]. Available: http://www.ifla.org/VI/7/pub/IFLAPublication-No119.pdf. 08.15.2012

[11] TG. Beamsley, "Securing digital image assets in museums and libraries: a risk management approach," Library Trends, Vol.48, No. 2, Fall 1999, pp. 359-378.

[12] The National Archives, Digital Preservation Guidance Note: Selecting File Formats for Long-Term Preservation, 2008, [Online]. Available:http://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf. 08.15.2012

[13] Wikipedia, "Open Archival Information System," *http://en.wikipedia.org/wiki/OAIS*. 08.15.2012

[14] A. Waugh, R. Wilkinson, B. Hills, and J. Dell'oro, "Preserving digital information forever," Proc. of the Fifth ACM Conference on ACM 2000 Digital Libraries, June, 2000, pp. 175-184.

[15] Andrew K. Pace, "Digital preservation: everything new is old again," Computer in Libraries, Vol. 20, Issue 2, Feb. 2000, pp. 55-57.

[16] Wiggins, Richard. "Digital Preservation Paradox & Promise," Library Journal, Vol. 126 Issue 7, Spring 2001, pp. 12-15.

[17] A. Muir, "Legal deposit of digital publications: a review of research and development activity," Proc. of the First ACM/IEEE-CS Joint Conference on Digital Libraries, Jun. 2001, pp. 165-173.

[18] Software Sustainability Institute, *http://software.ac.uk*. 08.15.2012

 ISBN: 978-1-61208-228-8

Technological Accessibility for People with an Intellectual Disability: A model for Organizational Support

Dany Lussier-Desrochers, Martin Caouette, and Marie-Ève Dupont
Department of psychoeducation
Université du Québec à Trois-Rivières
3351, boul. Des Forges, C.P. 500
Trois-Rivières, Québec, G9A 5H7
Dany.Lussier-Desrochers@uqtr.ca, Martin.Caouette1@uqtr.ca, Marie-Eve.Dupont@uqtr.ca

Abstract—**For a few years, technology has created an effective revolution in the field of intervention among people with intellectual disabilities. The use of such technology contributes to social inclusion and participation of the people in their community. Unfortunately, several obstacles are encountered by social services during the implementation of technologies among people with intellectual disabilities. A research with new ideas and new results was conducted with 11 managers in the field of social services to identify their needs and expectations as well as the difficulties they encounter in deploying technology during intervention. This research documents the mechanisms for social services and more specifically the elements that promotes an optimal access to technology. Results have allowed the creation of a management model supporting the application of technology in the field of social services which is called the Multidimensional Model of Organization's Support. The model is based on three fundamental and interrelated dimensions (management, clinical and technical) and 3 parameters (the user profile of a technological innovation, the type of technology and the life cycle of technologies). Preliminary results confirm the importance of providing different forms of support to clinicians in each dimension according to the identified parameters.**

Technological accessibility; intellectual disability; social services; organization

I. TECHNOLOGY AT THE SERVICE OF THE PERSON WITH DISABILITIES

Recent studies have shown that Information and communication technologies (ICT) have the potential to create a revolution in the different ways to assist people with an intellectual disability (ID) and to support them [1]. These technologies serve two major functions, to deputize for certain disabilities of the person and to promote the learning of new skills [2][3][4]. Currently, research results demonstrate the positive impacts of technologies when applied in intervention with people with ID. These technologies allow people to communicate, to have a valuable employment, to have access to independent apartments, to learn, to establish and maintain social contacts, to travel, etc. [4][5][6]. Their use promotes the inclusion of the people in their community and is a crucial determinant associated with social participation. Currently in Canada, these discoveries generate a lot of interest in the field of social services. The use of technology with people with intellectual disabilities is a real innovation for clinicians as far as it allows the development of practices that are more effective. A strong will, therefore, emerges in order to integrate these different technologies to intervention practices in intellectual disability so that more people can enjoy the benefits associated with the use of these technologies. However, the implementation of technology in the social field requires the consideration of a number of essential conditions.

A. Conditions associated with the implementation of technology in the field of social services

Although these technologies can create a revolution in specialized intervention, the social field intervention environments have only very little information on the conditions to be put in place in order to deploy adequately and successfully these different technological solutions. Many managers mistakenly believe that the only layoff technologies in their organization will be associated with an increased use by clinicians [7]. However, a successful and sustainable implementation of technologies requires the consideration of many factors (establishment of a vision on the role of technology, specific goals for implementation, precise evaluation of necessary resources (technical and professional), risk assessment, setting up an effective communication system, etc.) [8][9][10]. Moreover, it is essential to consider the technical characteristics of the technology to deploy, the user profiles and the life cycle of the technology. To effectively implement the technology in their organization, managers should review their management frameworks.

Unfortunately, managers do not have models to guide them in the deployment of technology in their organization. Therefore, the use of technology is achieved by a few clinicians on an individual basis and approaches are mostly "trial and error" procedure type. This approach is unfortunately not viable for a large-scale deployment of technologies in rehabilitation centers. The absence of a framework and guidelines for the management of

 ISBN: 978-1-61208-228-8

technologies used in the social services field is currently the main obstacle to implement the technology in these environments. The social service field should then develop management strategies to ensure efficiency while integrating these technologies without adding a financial burden to develop the practice in a context of fiscal restraint, taking into account the fragility of the process associated with the implementation of technological innovation. Research results shows that if those conditions are not taken into account, the technology will be implemented with a lot of difficulty and it will easily be untended by clinicians.

The paper will first present the qualitative research conducted with managers of rehabilitation centers. In a second section, the model developed by the research team is presented. The 3 dimensions and the 3 parameters are then described. The article concludes by presenting the next steps of the research project. Results and the model are briefly discussed in the conclusion.

II. PERCEPTIONS OF MANAGERS ON THE ROLE OF TECHNOLOGY IN THE FIELD OF SOCIAL SERVICES

In the next subsections we will present the objectives of the research, the method for data collection and results synthesis.

A. Research goal

The goal of this research is to develop and implement a business model guiding the implementation of technological innovation in Rehabilitation Centers for Intellectual Disabilities. This research is done in three phases. Phase 1 (2010-2011). Develop the management model based on interviews conducted with managers from Rehabilitation Centre for Intellectual Disabilities. Phase 2 (2012-2013). Implement the model through a pilot test in two Rehabilitation Centers for Intellectual Disabilities. Phase 3 (2012-2014). Documenting the implementation of the model and assess the impact of the technology use by the clinicians.

B. Method of Phase 1.

The paper presents the results of the first phase. This phase was carried out in two stages. First, interviews were conducted with managers. In a second step, a model was developed from the results obtained.

C. Interviews with managers

Our team conducted a survey with 11 executives from five Rehabilitation Centers for Intellectual Disabilities [11]. The purpose of these was to develop a managing model tailored to these fields of intervention. The research participants are managers who hold senior executive positions. On average, they were occupying this position for 2 years. However, they had in average 19.5 years of experience within their organization. The research used a convenience sampling. Telephone interviews of 25 to 40 minutes were conducted during summer of 2011 and winter of 2012. The interview had two major themes: the role of the technology in the organization and the conditions to be put in place to promote the implementation of technologies.

In order to fully explore the perceptions of senior executives, a qualitative approach was preferred. From the verbatim transcription, a thematic analysis was conducted. This form of analysis consisted at "proceed systematically in the identification, grouping, and, alternatively, discursive examination of topics in a corpus" [12]. The thematic analysis allowed us to identify relevant themes that were related to the objectives of our research and to document the importance of certain themes in our data corpus [12]. Analyses were performed using the software QSR N'Vivo 8 [13].

D. Main results obtained from the managers.

In general, senior managers have a positive perception of the potential of technology, but find that the implementation within their organization is still in its infancy. They identify problems in three dimensions: technological, clinical and management.

When considering the technological aspect, managers report that they are having inadequate knowledge of all the technology products available. Similarly, they are in lack of the equipment that is necessary (devices, incompatible computer systems). The choices to be made on the preferred technologies are also a problem, since they do not have accurate assessment of the potential of each technology to support clinical interventions. The clinical support for the staff on the technological dimension is also a major difficulty.

In terms of management, managers are finding that technology deployment within their organization is not done consistently and without a real structured action plan. Thus, they realize that they do not have a clear vision on the role of technology in their organization or specific goals of implementation. In addition, they operate in a context of fiscal restraint, which does not favor the investments in the technological dimension.

When considering the clinical aspect, managers realize that they do not have in their organization, clinical staff with the skills needed to implement and to use effectively various technologies. Moreover, there is only a small amount of training available to them on this dimension. Similarly, the social field has an insufficient number of professionals in order to explore possible avenues with technology.

Consequently, managers also report the need to structure the implementation of technologies in relation to these three dimensions. Furthermore, it appears necessary to define guidelines for the selection and purchase of technologies. Adequate support for clinical staff on the technological dimension is also majorly needed. Finally, managers want the technology implementation in their organization to be accompanied by an ongoing evaluation of the whole process.

E. Developing the management model

The first phase of the study made it possible to highlight some aspects to consider while implementing these technologies in those organizations. Consequently, these elements have permitted us to propose a multidimensional model of organizations for the implementation of supporting technologies. As mentioned earlier, managers had no

 ISBN: 978-1-61208-228-8

management structure in order to support technology deployment. The contribution of the research is to develop a customized model from an analysis of the situations faced by the professional in the social field services that works with developmental disabilities. The construction of this model was also the achieved by a literature review on the conditions to be deployed to ensure a successful implementation of technological innovation. The contribution of the model is the fact that it considers theories in all the dimensions such as marketing, management, information technology and intervention. The model is then located at the confluence of these disciplines. More precisely, it responds to the concerns of managers who recommend intervention in terms of clinical, technological and management aspect. The following sections present the dimensions of the model developed during the first phase.

III. MULTIDIMENSIONAL MODEL OF ORGANIZATION'S SUPPORT

The proposed model allows implementing all the conditions that will support a successful deployment of technology in an organization providing services to people with intellectual disabilities. This model takes into account three dimensions that needs to be consider in order to ensure a successful implementation of technologies named as: 1. The clinical dimension, 2. The management dimension and 3. The technical dimension. These three dimensions are constantly interacting with each other. The proposed model allows to coordinate the actions of all the key actors involved. It also formalizes the role of a techno-clinical counselor whose primary mandate is to coordinate the implementation of technologies taking into account these three dimensions. The purpose of the implementation process is to make the organizations autonomous when integrating different technologies to intervention practices in intellectual disability.

A. The management dimension

In terms of management, the model allows the organization to define a vision of the role of technology with people with disabilities and to implement the various processes necessary to maintain this vision. Thus, it allows to structure the ongoing integration of technological innovation within the organization. In order to achieve this goal, it promotes the establishment of a management structure that allows the evaluation, planning and reallocation of human, financial and technical resources that are necessary. The developed management tools allow managers to quickly assess the potential of technology, to make strategic choices while purchasing these technologies and to optimize the use of available computing resources.

B. The clinical dimension

Clinically, the model takes into account the contribution that technology can have on the quality of life and social participation of the people. It promotes the development of clinical intervention process that takes into consideration the technologies that are available. It has a structuring effect since it promotes a synergistic work with all the actors involved in order to develop and implement strategies to support the appropriation of technological innovations that consider the type of user, the type of technology, the life cycle of the technological innovation and time of appropriation. These support strategies can take the form of community of practice, communication tools, group of practice analysis, training, consultations, etc.. Thus, the model allows to set up boundaries so that the Organization can, independently and continuously, develop and adapt strategies of intervention according to the various types of technologies available.

C. The technical dimension

Technically, the model takes into account different technological parameters that must be considered in order to use technology in an effectively way. It also structures the ongoing and individualized support that must be offered to people who use technology in a clinical intervention perspective. This support guides managers during the budget planning and purchasing of computer equipment to support the intervention. The technical dimension also provides a constant update of the technology and a rapid replacement in case of broken hardware. These two elements are essential in order to prevent technical reasons that could lead to a cessation of the services provided to the people using technology in their daily lives which could have serious consequences for people with intellectual disabilities. Finally, the model formalizes procedures that facilitate testing, the use and loan of computer equipment to support the intervention.

D. The 3 parameters of the Multidimensional Model of Organization's Support

To effectively support actors working on these three levels, three parameters are considered: the type of technology, the user profiles and the life cycle of technologies. The user profile of a person influences his use of technology, but also his interest to make it accessible to others. Similarly, the type of technology modulates the energy that needs to be deployed in order to make the technology available to people. Finally, the life cycle of technology can target the times when the promotion of a technology must be achieved and when it should be abandoned. These parameters allows to modulate the type of support that will be offered: (communication, training, coaching the purpose and use, practice analysis, etc.).

Fig. 1 shows that in the context of the technologies deployment in an organization providing services to people with intellectual disabilities, the complexity lies in the establishment of a process taking into account not only the three parameters but also the dynamic process between them (adapted from Berkowitz et al. [14]). The establishment of a supporting targeted process is then the preferred solution. On one side, we should avoid putting into the hands of reluctant clinicians (minority appart) technologies that are radically changing their intervention process (long learning technologies); in those situations, technologies without a

 ISBN: 978-1-61208-228-8

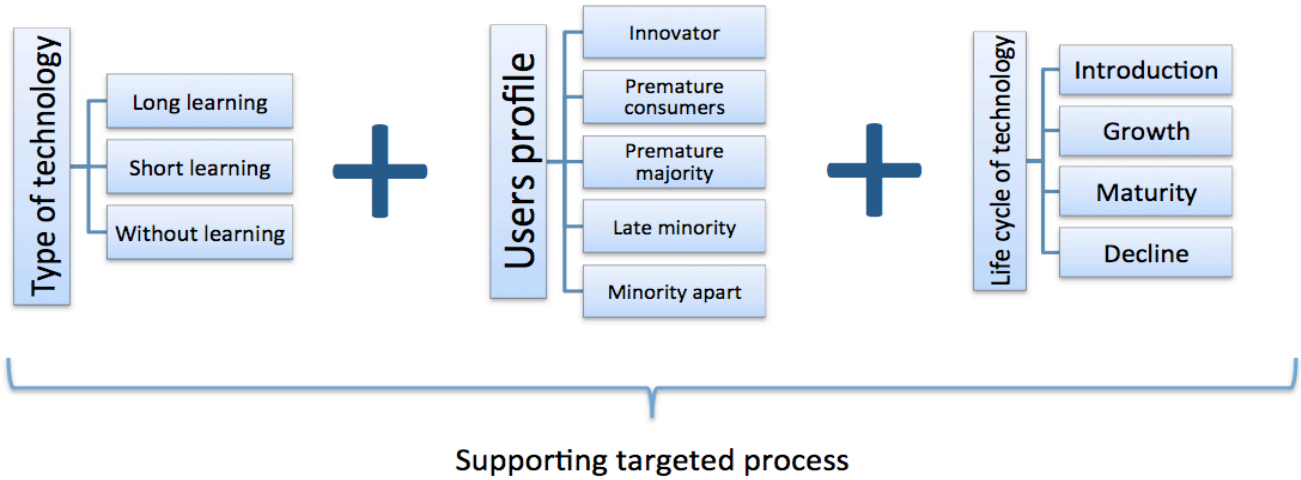

Figure 1. Multidimensional Model of Organization's Support

learning process will be favored (e.g., a software available online). Conversely, clinicians with an innovative profile will be attracted to more complex technologies (long learning) and have an interest in the technologies that requires a longer period of exploration or appropriation (eg iPad).

On the other hand, the field of technological innovation is an area of development at "high speed". The products are often replaced by new ones and life cycle of technological innovation is well known, as shown in the last component of the model, and it is essential to consider this fact. In the introductory phase, the technology reaches the market and the growth phase is then associated with the increasing presence of the product. During the maturity phase, the technology is used by many people who have integrated it into their lifestyle. In the decline phase, the technology is often replaced by a new product more powerful and more efficient. Schilling and Thérin (2006) [15] estimate the lifetime of a technological innovation (introduction to decline) to less than 5 years. In the field of social services, it means that the exploration phase and the training must take place in the growth phase. The application of the technology with the users must absolutely take place early in the maturity phase. This cycle needs to be performed again with a new technology during the declining phase of the first. The success is partly determined by a crucial strategic planning common to the managers and clinicians. Moreover, the life cycle of technology needs to be considered for people with intellectual disabilities, since it is necessary to establish specific procedures to ensure maximum stability against different types of intervention proposed.

E. Phases 2 and 3 (2012-2014)

The team is currently conducting Phase 2 by doing the implementation of the model in two rehabilitation centers. The validation of the management model is associated with a strategy of an action research in order to evaluate each step of the process (on-going evaluation process over two years to ensure a readjustment of the ways of supporting as needed). Indicators are then evaluated based on four pivotal concepts: 1) the level of ownership of the technological innovation through experimentation, 2) the impacts on intervention practices, 3) learning achievement regarding the technological innovation, and 4) the knowledge transfer and the impact of multipliers agents in the organization. At the end of the process, the tools for evaluating indicators will be reusable across all the organizations wanting to integrate the technological innovation. This ongoing evaluation of the process is carried out with clinicians and managers.

IV. CONCLUSION

In the last few years, the social work field had to face a massive influx of the implementation of supporting technologies used for diverse intervention with people with an ID. For these people, the impacts of these technologies on their lives are considerable. This research conducted with 11 managers demonstrated the importance of promoting the multidimensional implantation, taking into account three fundamental elements. This research has then allowed developing a model supporting the implementation of technologies in these environments. This model takes into account the clinical dimension (user profiles, the type and the life cycle of the technology), the technological dimension (support and up-to-date computing equipment) and the management dimension (assessment, planning, risk management and optimization). This model ensures that the organization remains master of the work process and it supports the key players in the development of new expertise. The model finally ensures a sustainable and successful implementation of technologies in support of the intervention. The model is a first step in the midst of social services allowing the efficiently implementation of the technology. This model will not only overcome the obstacles identified but it will also ensure an optimal use of technologies by people with intellectual disabilities. These

 ISBN: 978-1-61208-228-8

technologies will then have the effect of promoting social participation of the people and to reduce the negative impacts associated with digital exclusion of persons with disabilities.

REFERENCES

[1] D. Lussier-Desrochers, Y. Lachapelle, C. Consel, and D. Lavergne, “Utilisation de la domotique afin de promouvoir l’autodétermination et l’accès au milieu résidentiel pour les personnes présentant une déficience intellectuelle,” Recueil annuel d’ergothérapie, vol.3, 2010, pp. 9-20.

[2] D.K. Davies, S. E. Stock, and M. L. Wehmeyer, "Enhancing Independent Time-Management Skills of Individuals with Mental Retardation Using a Palmtop Personal Computer," Mental Retardation, vol. 40, Oct. 2002, pp. 358-365.

[3] S. E. Stock, D. K. Davies, M. L. Wehmeyer, and S. B. Palmer, "Evaluation of Cognitively Accessible Software to Increase Independent Access to Cellphone Technology for People with Intellectual Disability," Journal of Intellectual Disability Research, vol. 52, Dec. 2008, pp. 1155-1164, Doi: 10.1111/j.1365-2788.2008.01099.x

[4] Y. Lachapelle, D. Lussier-Desrochers, and H. Pigot, "Des TIC en soutien à l’autodétermination des personnes présentant une déficience intellectuelle," Revue québécoise de psychologie, vol. 28, 2007, pp. 111-123.

[5] M. M. L. Verdonschot, L. P. de Witte, E. Reichrath, W. H. E, Buntinx, and L. M. G. Curfs, "Impact of Environmental Factors on Community Participation of Persons with an Intellectual Disability: A Systematic Review," Journal of Intellectual Disability Research, vol. 53, Jan. 2009, pp. 54-64, doi: 10.1111/j.1365-2788.2008.01128.x.

[6] D. Zisimopoulos, J. Sigafoos, and G. Koutromanos, "Using Video Prompting and Constant Time Delay to Teach an Internet Search Basic Skill to Students with Intellectual Disabilities," Education and Training in Autism and Developmental Disabilities, vol. 46, Jun. 2011, pp. 238-250.

[7] S. Parsons, H. Daniels, J. Porter, and C. Robertson, "The use of ICT by adults with learning disabilities in day and residential services," British Journal of EducationalTechnology, vol. 37, Jan. 2006, pp. 31-44, doi:10.1111/j.1467-8535.2005.00516.x

[8] G. Corriveau, "Exceller dans la gestion de projet". Montréal, Qc : Collection Entreprendre, 2010.

[9] D. Lussier-Desrochers, M. Caouette, and M-E. Dupont, "Gestion de l’innovation technologique : défis et modalités d’accompagnement,″ Rapport de recherche adressé aux gestionnaires des CRDITED du Québec, Université du Québec à Trois-Rivières. Repéré à http://www.chairetsa.ca/journeetsa/Implantation.html , 2011.

[10] M-È. Dupont, "Identification des conditions de succès liés à l'implantation et à la périnité d'un site internet spécifiquement adapté aux personnes qui présentent une déficience intellectuelle" (Mémoire de maîtrise inédit). Université du Québec à Trois-Rivières, QC, 2012.

[11] D. Lussier-Desrochers, and M. Caouette, "Pourquoi la technologie en soutien à l’intervention ne s’implante-t-elle pas plus rapidement dans les milieux d’intervention?" Revue du CNRIS, (submitted).

[12] P. Paillé, and A. Mucchielli, "L'analyse qualitative en sciences humaines et sociales" Paris: Armand Coli, 2008, pp.162.

[13] QSR International Pty Ltd. NVivo 8. Australia. 2008. ABN: 47 006 357 213

[14] E.N. Berkowitz, F.G. Crane, R.A. Kerin, W. Rudelius, D. Pettigrew, S. Gauvin and W. Menvielle. Le marketing. Montréal, QC: Chenelière McGraw-Hill, 2007.

[15] M. Schilling and F. Thérin, "Gestion de l’innovation technologique". Paris : Maxima, 2006

 ISBN: 978-1-61208-228-8

Use of ICT in Finnish Social Welfare

Situation in 2011

Jarmo Kärki, Maarit Laaksonen, Hannele Hyppönen
National Institute for Health and Welfare
Information Department
Helsinki, Finland
jarmo.karki@thl.fi, maarit.laaksonen.2@thl.fi, hannele.hypponen@thl.fi

***Abstract*— National level monitoring of adoption and use of information systems is becoming common practice, providing one set of indicators for success of e-health policies. This is not the case in social welfare, in spite of advancing national e-welfare policies. This article presents results of a national web-based survey of availability and use of electronic social services and social welfare client information systems in Finland. The survey conducted in 2011 was targeted at managers or IT-leaders of public and private service provider organizations. The results show that the majority of social welfare service providers in municipalities have a client information system (CIS). A half of the public social welfare organizations use mobile systems for some purposes. Private service providers use CIS's somewhat less than local authorities. Small private enterprises generally do not have CIS's. Not all social welfare employees have a personal workstation. Information exchange between organizations is limited to read-only. There is little exchange of information between information systems. Internet access is more commonly available. Providers of social welfare services generally have a website with information on their operations. Public service providers are more likely to have their own IT personnel than private service providers. The percentage of CIS investment costs out of total ICT costs varies greatly from one organization to another. Local authorities invest more in ICT than private service providers. The development of e-welfare services has been rather slow over the past decade. Results will be used in development of national information services for social care in Finland.**

Keywords-social care; social services; private social welfare service providers; ICT; electronic information management; client information systems.

I. INTRODUCTION

E-government agendas, regulatory bodies and Information and Communication Technology companies are participating in shaping what has been called an 'electronic turn' in social work and social care [1]. There is an increasing interest in the European Union, the World Health Organization and The Organisation for Economic Co-operation and Development as well as national authorities in monitoring the success of this turn [2][3], but the focus has been largely in e-health, not in e-welfare [4][5]. A recent literature review [6] searched five databases, such as JStor, Annual Reviews, Ebsco, PubMed and Nelli, and three journals, as Communications in Information Literacy, Journal of Social Work and British Journal of Social Work refereed electronic journals. The search was conducted combining the following search terms: social* AND informatics, social* AND information technology, social* AND information systems and social* AND knowledge management. The review revealed lack of studies of e-welfare adoption. According to the review, the studies focus mainly usability or use of applications in social services. Client's perspective or the social work processes have been ignored. Situation is the same in Finland: e-welfare execution and the use of ICT in social welfare has been monitored previously in 2001 [7]. Internationally, similar studies are rare [8].

However, e-government and e-welfare policies have advanced remarkably during the 10-year period in Finland: an act on Information Management Governance in Public Administration entered into force in 2011 [9]. An ambitious purpose of the law is to lay down provisions on information management governance in public administration and to ensure the interoperability of information systems. The Finnish Ministry of Finance is devising an overall enterprise architecture for public administration. The authorities in central and local government, as well as in social welfare, each plans and specifies their enterprise architecture in accordance the overall architecture. As a part of public services, the horizontal interoperability of client information systems in social welfare as well as the vertical interoperability of governmental information systems, needs to be improved. The client information must be available for the social welfare authorities wherever and whenever it is needed. The citizens have a right to seamless and user-friendly e-services and e-practices in social welfare. The aim of citizens' e-services is to increase service availability where and when needed and empower clients in their own care.

Many academic and practical models have been proposed to understand the maturity of digital government or e-government [10]. Often the stages of e-services, e-democracy or other e-practices development are described in these models. Initial stages concentrate on the availability of public information and on the possibility of handling administrative processes online. Later stages are characterized by for example vertical and horizontal system integrations, a homogenous network of public services and data mobility and sharing across public and private services.

 ISBN: 978-1-61208-228-8

The models and the assessment indicators have been criticized to be predominantly technical overshadowing crucial social aspects, such as power, interest groups, conflict and values [11].

Although interesting research of electronic recording in social welfare and social work practices are published [12][13][14][15][16][17], the need for the advanced research methods of social care informatics has been widely recognized [18][19][20][21]. Also a Declaration for action has been presented by twenty-three actors with a multidisciplinary background of research from 15 different European countries [5]. The Declaration of action is a significant effort to define social care informatics more rigorously as a field of research and as a domain of research investigations.

The purpose of this article is to present results of a national survey of electronic social services and social welfare client information systems currently available for key social services in Finland [22]. The questions that are studied in this paper are:

(1) What kind of e-services has been offered for citizens?

(2) What kind of client information systems and other professional e-tools the social welfare employees use?

(3) To what extent is information exchange between organizations possible in the social welfare services?

(4) How does the data management and resources used to ICT look like in social welfare sector?

The survey was conducted as part of the Government Programme for public e-services for citizens (SADe programme) [2] funded by the Ministry of Finance, where electronic service systems are developed for different public services, including the social welfare and health care sector. The survey was commissioned by the National Institute for Health and Welfare and conducted by Market-Visio Oy. The data concerning the availability and use of ICT in health care were collected simultaneously and the results were published in a separate report.

First the methods and data of the survey are presented. Then the main results of the survey are reported and two figures presented. Finally the conclusions, future works and implications of the study are presented in Section 4.

II. METHODS AND DATA

Data collection was carried out as semi-structured Webropol-questionnaire at the beginning of 2011. The questionnaire was modified from an instrument that has been used for mapping health care IT diffusion and use biannually since 2003, in order to maintain comparability of the data between sectors. The target group of the survey were all the chief information officers of the Finnish local government organizations organising social services, and private social service providers in Finland.

There were 336 municipalities and 35 municipal federations in Finland at the time of the survey. In 2010, there were 2 922 private organizations, associations or companies providing social services, which delivers almost a third of all social services in Finland. All organizations of the target group could not be reached.

In the social services part of the survey, 457 organizations providing social welfare services responded: 69 local authorities, 13 municipal federations (or similar bodies) and 373 private enterprises, associations or non-governmental organizations. Response rate (23 %) remained low despite the best efforts to acquire additional responses. However, organizations of all sizes responded to the survey. Geographically the respondents represent whole Finland. As far as public social services go, the responses cover some 63% of the population of Finland. The data includes all the statutory social welfare service tasks, such as child protection, social work, disability services, substance abuse services and services for older people. The responses from private service providers were mainly from organizations providing services for the elderly such as home services, housing services and institutional care.

The limited data do not allow for meaningful statistical inference. Therefore, this article focuses on the descriptive statistical analysis. The results are interpreted as a sample of the target group. The results are presented graphically by describing the values of the distributions and cross-tabulating.

III. RESULTS

The survey asked what kind of public electronic services the respondents organize for citizens. Providers of social welfare services generally have a website with information on their operations. Half of the local authorities and one fourth of the private service providers had a facility for online feedback. One in four of the local authorities responding reported that they offered online services for citizens such as applications for daycare places or income support. In cases where citizens were offered a facility for e-transactions, about half of the clients choose to manage their affairs with the service provider in some other way. Anyhow eSocial care is about to come: online services are being developed in several projects around the country. Overall architecture for e-services was also asked by the local actors.

The survey demonstrates that the majority of providers of social welfare services in local government have a client information system. Two applications dominate the market, as shown in Figure 1. One-fifth of the public social welfare organizations use more than one CIS's, having different CIS for different services.

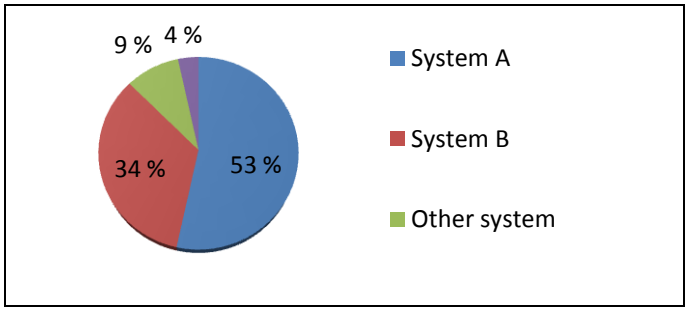

Figure 1. The percentages of different client information systems used by public social service providers (% of respondents) n=89.

 ISBN: 978-1-61208-228-8

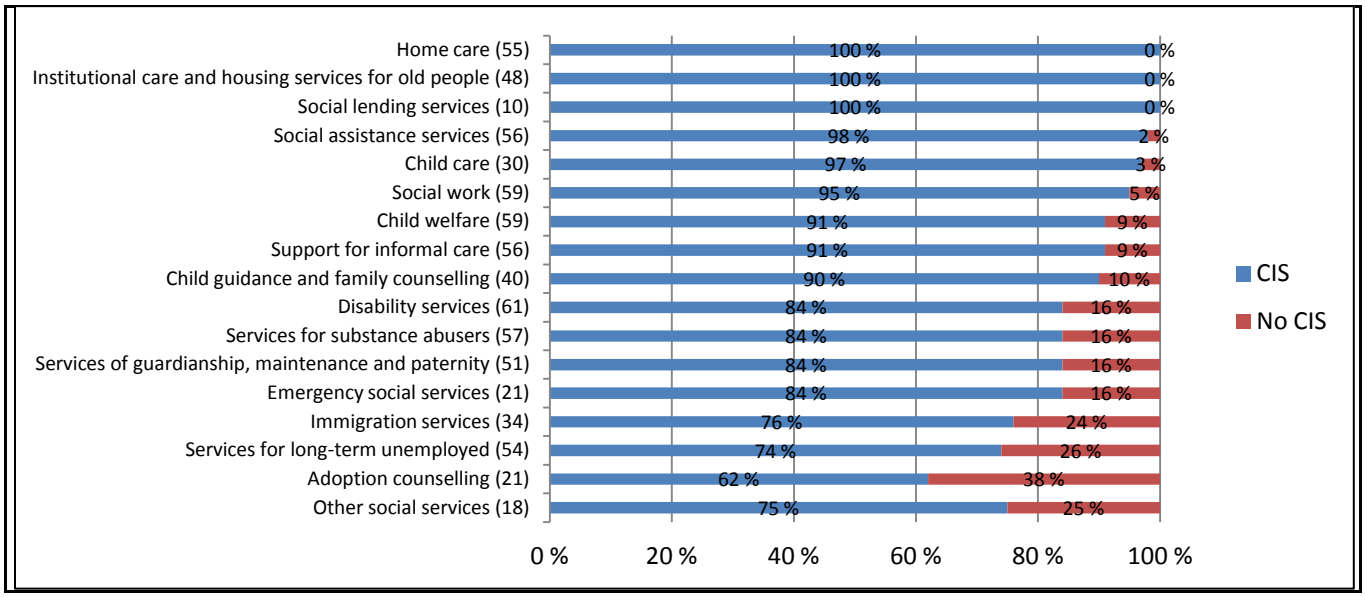

Figure 2. The percentages of the different public social services using client information systems. The data of the public social service providers (% of respondents). The number of respondents is presented in the parenthesis after the name of the social service.

However, the possibilities for an employee in municipal social services to use these client information systems depends on the particular service and access rights of employees; not all of the branches of the social services have a client information system. For example a quarter of those working in adoption counseling or immigration services are missing the opportunity to record on client information systems. The percentages of the different public social services using or not using CIS's are presented in Figure 2.

Private service providers use client information systems somewhat less than local authorities, although there is great variation. Small private enterprises generally do not have information systems.

The contents, like classifications or data structures, of CIS's vary depending on software and organization. Requirements of an annual collection of national statistics are so far the only unifying element between the different systems.

The surveys reveals that local authorities commonly use national classifications for compiling the annual statistics for the National Institute for Health and Welfare, such as the statistical classifications in social services institutional and housing services notifications, child welfare services statistics classifications, and child support and maintenance statistics classifications. Private service providers use only the statistical classifications of care notifications.

The survey shows that information exchange between organizations is so far limited to read-only, and there is little exchange of information between information systems. Only one of five public service providers reported that digital exchange of information is possible. The results show that municipal social services have quite good access to the query system SOKY of the Social Insurance Institution (Kela) and the population register system of the Population Register Centre (VRK-database). The personal information of these systems is used by social workers in particular when granting social assistance or assessing the clients' needs in the social work. Without electronic access to this data, the same information is acquired by telephone or in a printed form delivered by the client. Nearly two of five public service providers have organized an access to an electronic patient record of primary health care. Access to other information systems is considerably more rare.

Also, access to client information outside organizations is strictly limited to designated employees specially those working in social assistance, social work and child protection. In private organizations, employees rarely have access to outside information systems, and information exchange between systems is extremely rare.

Not all social welfare employees in the public or private sector have a personal workstation. Internet access is more commonly available. Almost half of the public social welfare organizations use mobile client information systems for some purposes, and the other half does not have the similar possibility. Private service providers, by contrast, have virtually no mobile client information systems. Social services professionals commonly use their own user identifier and password to identify themselves. Only a few of the public-sector organizations responding are using an official e-transaction card, a health care certificate card or other identifier device.

Public service providers are more likely to have their own IT personnel than private service providers (68 % vs. 38 %). Similarly, public service providers are more likely to have an information management strategy, an electronic archive plan, client documentation instructions for employees or data protection or information security instructions than private service providers. Altogether, there are significant deficiencies in social care information management in Finland.

The survey indicates that the online professional tools most frequently used by social services employees are the

 ISBN: 978-1-61208-228-8

Sosiaaliportti online portal [23], which publishes professional information for Finnish social workers, and the Intranets of their respective organizations. One in five public organizations and one in ten private ones make use of online learning environments.

The percentage of client information system investment costs out of total ICT costs varies greatly from one organization to another. Local authorities would seem to invest more in ICT than private service providers. Many respondents reported a percentage of less than 10 %, but some more than 50 %. ICT costs of the public service providers were total of about 40-50 million euro in 2010 based on an estimation derived from the results. It's about five percent of all ICT costs of municipalities in Finland. The majority of the respondents estimated that their total ICT costs would increase between 2011 and 2013.

IV. CONCLUSION AND FUTURE WORK

Limitations of the study include poor response rate, even if the sample was nationally representative. The data doesn't allow meaningful statistical inference and in order to illustrate the evolution of the ICT use. However, results can be compared to the situation in 2001 [7] and careful suggestions be made in light of key e-government and e-welfare policy goals to indicate most acute development needs.

There are websites and some facilities for online feedback, but well-designed online services offered for the citizens are still rare. Interestingly as many as 45 percent of the public social service providers announced to offer guidance via e-mail in 2001 and improving the e-services were mentioned as a key development area at that time. In the light of the findings, very rapid development hasn't taken place for the last ten years. Information about services is fragmented and access to them limited to certain times and places, with poor access to information required for applications. Regarding e-services for citizens as a significant e-welfare policy goal, the maturity level of the social welfare system seem yet to be low: the technology used is rather undeveloped and the distribution of the e-services has been limited to only a few social services. Social services are not yet available where and when needed, neither do they empower clients in their own care. More structured development of e-welfare services for citizens in SADe-program could start from a simple integrated service directory and electronic application of services with integration to national SOKY and VRK -databases.

Most of the client data can be recorded on the stationary CIS's engineered by two software houses. The logical level, such as classifications and data structures, of the CIS's is heterogeneous, which makes the information exchange between organizations difficult. Over the last decade the number of the municipalities, in which more than one CIS's are used, have increased. Despite of this software differentiation, there are yet some social service organizations, often with a limited number of clients, without any CIS's. One of the main e-welfare policy goals, the horizontal interoperability of client information systems in social welfare, is not yet possible. In order to achieve the goal, it seems necessary to plan, specify and implement the overall information architecture also in social welfare.

Social welfare professionals have access to the Internet and they log in the applications with the user ID's of their own, but only a limited group of the municipal social welfare employees have an access to the two governmental data systems. However, compared to the situation in 2001, the possibility to use information systems outside the home organization has improved considerably. Particularly, an option to use the digital information recorded in the primary health care services in elderly and home care seems to easier nowadays than 10 years ago. At least in some services, the progress has been made in vertical interoperability, when considering the availability of data for professional use in social welfare. However, there are still plenty of challenges regarding usability and advanced technological solutions.

Altogether, several millions of euro are laid every year out to the ICT of the social welfare in Finland, but the foundations, such as an enterprise architecture and the information management, need to be better organized. Even if there are CIS's and access to some information systems, the technology is only as good as the management and organizational systems, within which it is placed [24].

In Finland, which is said to belong to the Scandinavian and universalistic model of welfare regimes [25], citizens' rights to social protection are rooted in the Constitution. The municipalities have a responsibility to organize social welfare services for the citizens. The services are mainly produced by public institutions, too. As a result of this, public service providers have larger number of clients and more financial and human resources to invest in ICT than private service providers, which explains most of the differences of the results between public and private sector.

The Finnish e-welfare services are probably placed mainly on the initial stages of the maturity assessment models of e-government. Despite the Finnish high level Government information society strategies, legislation and Action Programmes [26], the results of this study shows that the basic infrastructure and the management of e-social welfare in Finland is tenuous and deficient. From the point of view of the social welfare professionals and of their clients speaking about "network society" [27] or information society ring hollow.

The results can be leveraged in further specifying the needs, potential and capacity for implementing nationwide online client service functions in social welfare and health care and for implementing nationwide social welfare information system services. The results may also be used when designing local and regional solutions.

This study paints a partially fragmented and patchy picture of e-welfare implementation in Finland, with more ICT support for certain social work areas than others. With e-government strategies being increasingly implemented, importance of monitoring the progress and impacts of them is growing for evidence based management of e-welfare services. The emerging software architecture and the greater use of centrally devised e-assessment templates attempt to map the structure of social work and construct new "workflows". It is important to ensure that this IT-mediated

 ISBN: 978-1-61208-228-8

(re)construction will happen in a transparent way, where feedback is used to steer further development. Feedback needs to be collected not only to monitor speed of the policy implementation, but also the acceptance and added value to the clients, workers and service providers [1][8][11][13][21].

The future development needs for this survey lie in closer integration with the e-health survey to monitor integrated care, and in extending the data collection to usability, user satisfaction and cost and benefits of the ICT-mediated services. So far, there are no systematic mechanisms to measure the customer satisfaction or the accuracy of the social welfare service system neither in municipalities nor in private social service organizations in Finland.

REFERENCES

[1] P.N. Garrett, "Social work's electronic turn: notes on the deployment of information and communication technologies in social work with children and families," Critical Social Policy 25, 2005, pp. 529 - 553.

[2] i2010 High Level Group. Benchmarking Digital Europe 2011-2015 a conceptual framework. Issue No: 27, October 2009. http://ec.europa.eu/information_society/eeurope/i2010/docs/benchmarking/ benchmarking_digital_europe_2011-2015.pdf. (10.1.2012)

[3] Global eHealth – Measuring Outcomes: Why, What and How. Bellagio, Italy, July 13-August 8 2008. http://www.ehealth-connection.org/files/conf-materials/Global%20eHealth%20-%20Measuring%20Outcomes_0.pdf. (10.1.2012)

[4] M. Rigby, P. Hill, and A. Allman, "Ensuring Trust and Governance in E-social Care," in Collaboration and the Knowledge Economy: Issues, Applications, Case Studies, Vol 5, Part 1, P. Cunningham, M. Cunningham, Eds. IOS Press, 2008, pp. 22-28,

[5] M. Rigby, P. Hill, S. Koch, and D. Keeling, "Social care informatics as an essential part of holistic health care: A call for action," International Journal of Medical Informatics 80, 2011, 544-554.

[6] M. Laaksonen, Social Work Informatics: Integrated literature review [Sosiaalityön tiedonhallinnan tutkimus: Integroitu kirjallisuuskatsaus], Master's thesis, Kuopio: University of Eastern Finland, 2010.

[7] K. Hartikainen, S. Kuusisto-Niemi, and E. Lehtonen, Survey of information systems in Finnish health and social care in 2001 [Sosiaali- ja terveydenhuollon tietojärjestelmäkartoitus 2001], Osaavien keskusten julkaisuja 1/2002, Helsinki, 2002.

[8] Department of Health. National Electronic Social Care Record Survey Report, 2006/07, England. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_075530.pdf. (26.7.2012)

[9] Act on Information Management Governance in Public Administration (634/2011), The Ministry of Finance in Finland, 2011. http://www.vm.fi/vm/en/04_publications_and_documents/03_documents/20110902ActonI/Tietohallintolaki_englanniksi.pdf (6.8.2012)

[10] S. Assad, I. Boughzala, and I. Boydens, "Back to Practice, A Decade of Research in E-Government," in Practical Studies in E-Government: Best Practices from Around the World, S. Assad et al., eds. New York: Springer, pp. 1-12, 2011.

[11] M. Simões, "Sociological Reflections on E-government," IARIA, ICDS 2012: The Sixth International Conference on Digital Society, 2012, pp. 29-34. http://www.thinkmind.org/index.php?view=article&articleid=icds_2012_2_20_10116 (3.8.2012)

[12] F. Reamer, "Documentation in Social Work: Evolving Ethical and Risk-Management Standards," Social work, vol. 50, pp. 325-334, 2005.

[13] J. Kagle, "Record Keeping: Directions for the 1990's," Social Work, vol 38, pp. 190-196, March 1993.

[14] I. Shaw, M. Bell, I. Sinclair, P. Sloper, W. Mitchell, P. Dyson, J. Clayden, and J. Rafferty, "An exemplary scheme? An evaluation of the integrated children's system," British Journal of Social Work, vol. 39, pp. 613–626, 2006.

[15] N. Ames, "Social work recording: A new look at an old issue," Journal of Social Work Education, vol. 35, pp. 227-237, 1999.

[16] J.-M. Räsänen, "Producing Norm Talk of Fact-based Case Recording in Interviews With Emergency Social Workers," Qualitative Social Work, vol. 11, pp. 6–22, January 2012.

[17] B. Pahwa, "Technology and School Social Work Services: Introducing Technology in an Alternative School," Journal of Technology in Human Services, vol. 21, No. 1/2, pp. 139-160, 2003.

[18] H. Mouratidis, G. Manson, and I. Philp, "Analysis and design of the eSAP: An integrated health and social care information system," Health Informatics Journal, vol. 9, No: 2, pp. 89–92, June 2003.

[19] M. Rigby, P. Hill, S. Koch, and J. Kärki, "Social care informatics – the missing partner in eHealth," in Medical Informatics in a United and Healthy Europe, Proceedings of MIE 2009, K.-P. Adlassnig et al., Eds. Amsterdam: IOS Press, 2009, pp. 277–281.

[20] D. Parker-Oliver and G. Demiris, "Social Work Informatics: A New Specialty," Social Work, vol. 51, No: 2, pp. 127-134, 2006.

[21] P. Hill, A. Allman, M. Rigby, and R. Wilson, "Social Care Informatics: Beginning to Face up to the Reality of the Mixed Services Economy?" in Collaboration and the Knowledge Economy: Issues, Applications, Case Studies, Vol 5, Part 1, P. Cunningham, M. Cunningham, Eds. IOS Press, 2008, pp. 14-21.

[22] J. Kärki, M. Laaksonen, and H. Hyppönen, Use of technology in Finnish social welfare in 2011 [Tieto- ja viestintäteknologian käyttö sosiaalihuollossa vuonna 2011], Raportti 2/2012, Helsinki: Terveyden ja hyvinvoinnin laitos (THL), 2012.

[23] http://www.sosiaaliportti.fi/en-GB/(10.1.2012)

[24] T. Carrilio, "Management Information Systems: Why Are They Underutilized in the Social Services?" Administration in Social Work, vol. 29, No: 2, pp. 43-61, 2005.

[25] G. Esping-Andersen, "The Sustainability of Welfare States into the Twenty-First Century," International Journal of Health Services, vol. 30, No: 1, pp. 1-12, 2000.

[26] The National Knowledge Society Strategy 2007-2015. Finland, 26.9.2006, http://www.tietoyhteiskuntaohjelma.fi/esittely/en_GB/introduction/_files/76222690188788831/default/Strategia_englanti_181006final.pdf (6.8.2012)

[27] M. Castells and P. Himanen, The Information Society and the Welfare State: The Finnish Model, New York: Oxford University Press Inc, 2002.

 ISBN: 978-1-61208-228-8

Distribution of Municipal Services Access Points over a Territory in a Scenario with Restricted Information

Mariuxi Montes*, Miguel Angel Peña†, Susana Muñoz Hernández‡
Facultad de Informática
Universidad Politécnica de Madrid, Madrid, Spain
* Email: maria.montes.chunga@alumnos.upm.es
†Email: m.pena@upm.es
‡Email: susana@fi.upm.es

***Abstract*—In developing countries we face not only technical issues, but also the lack of basic data that we are obly to obtain by using alternative paths. In the context of a local government on a developing country, this work shows two results. The first one is the procedure to obtain or deduce the needed information from alternative data sources. The second one is how to use genetic algorithms (whose input is obtained from the first result) to configure a distribution of access points for eGovernment services.**

***Index Terms*—eGovernment, Genetic Algorithm, developing countries, access to eGovernment services**

I. INTRODUCTION

This paper proposes a solution to a problem faced when planning eGovernment services for citizens of a municipality with characteristics such as: high rural population, very disperse towns, low Internet penetration and low incomes. First, the topic of eGovernment is discussed. Next, the problem facing the Santa Elena municipality is detailed. Finally, genetic algorithms are introduced as a tool to find a solution to this problem.

A. eGovernment

To successfully implement eGovernment initiatives, an evaluation of involved subsystems such as social, technical and environmental ones is needed [1]. eGovernment is not only about delivery, it is also about a cultural change in how citizens relate with the government [2].

In developing countries where access to Internet is restricted, delivery of eGovernment services raises an important question: how can it be made accessible to citizens?

B. The case of Santa Elena Municipality - Ecuador

The Municipality of Santa Elena is on the coast of Ecuador (see Fig 1.). It has a population of 144,076 where almost 72% of it is rural. These rural zones are very dispersed and every concentration of people is called a ”comuna”. There is only one ”large” city, the municipal capital, and encompasses a bit more than the 30% of the total population.

According to data from the Integrated System of Social Indicators of Ecuador - 2010, 76% of the population of the province live in poverty (as measured by the satisfaction of basic needs including housing, health, education and employment). 63% of the population has completed basic education and 9% have obtained higher education.

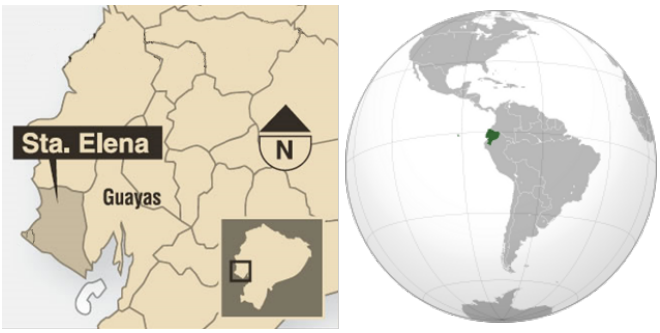

Fig. 1. location of Santa Elena Municipality. Figure based on graphics obtained from www.hoy.com.ec and www.wikipedia.com

There are 59 comunas distributed throughout the geographic territory according to the ”Federación de Comunas de Santa Elena”. However, lower concentrations of people still exist. These are known as ”recintos”, but any accessible formal statistics about the number of recintos in the territory are unavailable.

The majority of comunas are located on the coastline and they have good road access with the capital. However, the inner comunas have major access problems between them as well as with the capital due to a poor secondary road network. In addition, the connections between inner comunas are inadequate because of a lack of maintenance of existing roads.

Internet penetration in Ecuador, according to data from December 2008, is 12.3%, compared to the South American average of 30%. This ranks Ecuador 8/10 in South America [3]. The ”Superintendencia de Telecomunicaciones” [4] estimates that 37,46% of the total population in Ecuador are Internet users. The province of Santa Elena represents 0,75% of this number. These percentages are growing every year, but

 ISBN: 978-1-61208-228-8

these numbers are still too low.

Due to these statistics, it is necessary to take into account this low penetration when offering eGovernment services in this territory. One option for facing this issue is to use the strategy shown on [5], which is to use service access points (SAPs) to assure that citizens have access to eGovernment services. SAPs will manage the problem of low Internet penetration and make access to eGovernment services more readily available to citizens in this area. These SAPs would be geographically distributed over the territory in a way that efficiently covers the entire Municipality. This option allows the best coverage for the majority of the population instead of having all services centralized in the municipal capital.

How can this distribution be achieved across this territory? Distribution refers to the quantity and the location of the SAPs throughout the area. This is the question that guides this study. To answer this question, a Genetic Algorithm is used to show an adequate distribution.

C. Genetic Algorithm

Genetic algorithms are adaptive methods and can be used to solve search and optimization problems. John Holland introduced the term ”Genetic Algorithm” in his publication ”Adaptation in Natural and Artificial Systems” [6]. These algorithms are well defined in texts like ”Genetic Algorithms in Search, Optimization & Machine Learning” [7] or ”Modern Heuristic Techniques for Combinatorial Problems” [8].

Genetic algorithms try to mimic the genetic process of living organisms, which evolve generation after generation under the principles of natural selection and survival of the fittest based on Darwin’s Theory of Evolution. These algorithms create the most adaptable solutions for the proposed problems.

II. Material and Methods

This section will show how data was obtained, with an emphasis on the difficulty of obtaining it and the process used to acquire it. The fitness equation used to get a response to the question will also be explained.

A. Obtaining the Data

To address this task, we need some basic information. In the context of a developing country, this is the first and major problem to face because it is not always possible to obtain the necessary data.

The ”Instituto Nacional de Estadisticas y Censos” (INEC) is the institution in charge of maintaining and offering information about the demographics of Ecuador. INEC publishes this information in detailed level, including: national, provincial, cantonal and parochial. For the purposes of this study, a lower level, the zonal level, is needed to know the population of every comuna and recinto. To obtain the information needed for this level, the census database of the province was downloaded. Although this database detailed the population at the zonal level, a problem was encountered with this information. Because every zone is identified through census codes and not by name, it was impossible to obtain the census codes for all

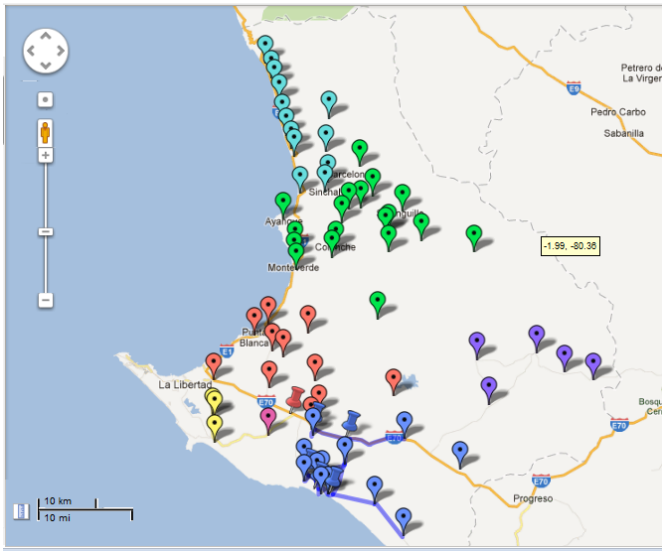

Fig. 2. Distribution of comunas through Santa Elena Municipality - Ecuador

comunas and recintos. Thus, only those comunas and recintos, which were identifiable through census codes were used.

Figure 2 shows an image of the Municipality and distribution of the comunas and recintos throughout it. On this map we can see colored points where each color represents a specific parish. There are 7 parishes that encompass a total of the 56 comunas and recintos selected to carry out this study. This figure only represents the comunas and recintos selected and not all of the existing comunas.

To measure the distances between points (comunas, recintos and crossroads), maps from Google Maps were used because of the layout of the main roads. However, Google Maps does not lay out the secondary roads and thus the use of maps in PDF format of every parish was necessary. All of these maps were downloaded from the INEC webpage. This data was obtained manually, using the measurement tools provided by Google Maps as well as by Adobe Acrobat, to measure distances between comunas and recintos and between crossroads that establish different roads to reach a determined comuna or recinto.

The identified communication roads are from various levels: (1) There is one main road connecting comunas settled on the coastline; this road is part of the E1 national highway. This road is in excellent condition and has continuous maintenance work. It is usable all year at a maximum velocity of 100Km/h. (2) There is a part of this E1 national highway that connects other groups of coastal comunas but its maximum velocity is reduced by 10 Km/h, resulting in a maximum velocity of 90 Km/h all year. (3) There are secondary roads, which connect inner comunas and recintos that are not settled on the coastline. These roads are usable all year, although they do not have good maintenance. These roads are narrower than the roads mentioned above and thus their maximum velocity is 60 Km/h usable all year. (4) Finally, there are ”summer roads” that are usable in dry season but during rain season it is difficult to

 ISBN: 978-1-61208-228-8

drive over these roads. These roads have a maximum velocity of 30 Km/h. The Figure 3 details the kind of maps used to measure distances between points and the levels (types) of existing communication roads between them.

An approximation of the time needed to communicate between each comuna and recinto was made. This was done considering the distance between comunas and recintos and the road level (maximum velocity) that communicates them. To make this approximation, the Floyd-Warshall algorithm [9] was used. This process is use to assure the lowest communication time between two towns amongst all the possible communication options available.

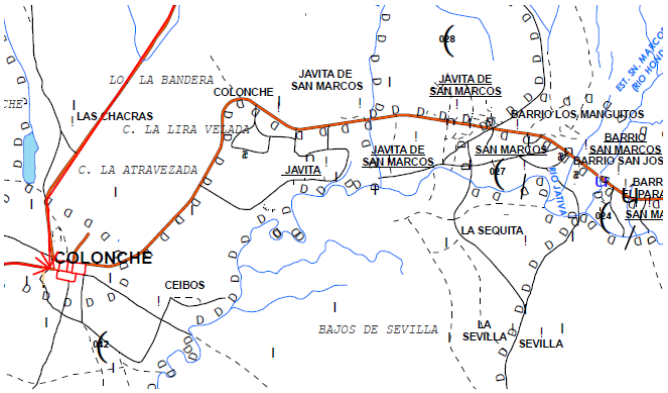

Fig. 3. Map, which details roads that connect comunas. Santa Elena - Ecuador

There is useful information about bus routes and frequencies that was unavailable.

B. Personalizing the Genetic Algorithm

It is necessary to establish a representation of the problem using the data about th number of selected comunas. Every individual has 56 genes (l=56), one for each comuna or recinto that would be able to host one SAP. If the gen is OFF, it is equivalent to a specific comuna or recinto with no SAP, and when the gen is ON, it is equivalent to a specific comuna or recinto with a host SAP. One possible configuration of an individual chromosome would be:

11000111000000000000000000000
00001100010000001111000000000

Then, the objective function is defined. To define it, information on the population of comunas or recintos and displacement times is considered. This function has to achieve a balance between citizens displacement time to the closest SAP and the cost of implementing a SAP.

An adjustment has been made to work with two distinct units: the population that moves and the cost of having more access points. The function punishes these two possibilities: moving the entire population to the main town or putting an access point in each town.

A weighting was made, giving more weight (0.6) to SAP implementation cost and a lower weight (0.4) to the citizens' displacement time.

The objective function is:

$$0,4\left(\frac{\sum_{i=0}^{l}(population_i * distance(i, CloserSAP))l}{\sum_{j=0}^{l}(population_j * distance(j, MainCity))}\right) + 0,6(TotalSAP) \quad (1)$$

where $l = 55$ (total number of comunas and recintos)

- *Initial population* would be $\lambda = 2l$, as suggested by Alander in [10]
- *Stop* or the maximum number of iterations is configured on 200
- *Crossover method* selected was one point cross.
- *Replacement method* is partial change, where a part (2 individuals) of the parents is maintained and the other part is of the children.
- *Mutation method* is aleatory mutation.
- *Mutation probability* is $1/l$ like De Jong suggest in [11].

III. Results

The first execution time took 3 seconds, and after running the algorithm, it gave a solution with a fitness value of 7.47 represented like:

00000000000000000000000100010
00000000000000000010010000100

Five comunas were selected to host a SAP from the total of 56. Translating this numerical representation to a map, the figure 4 shows a graphic where the selected comunas are highlighted.

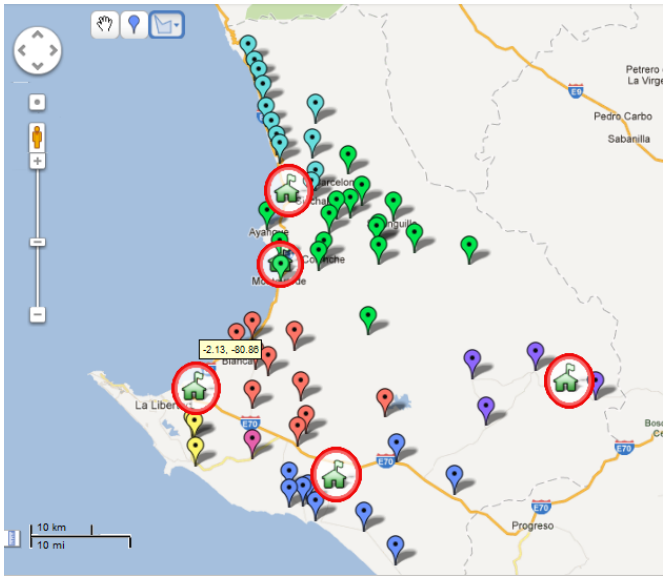

Fig. 4. Selected comunas to host a SAP

After this first execution, the algorithm was run more times, obtaining similar fitness values. Figure 5 shows that the solutions given by the genetic algorithm are improved with every iteration (population).

 ISBN: 978-1-61208-228-8

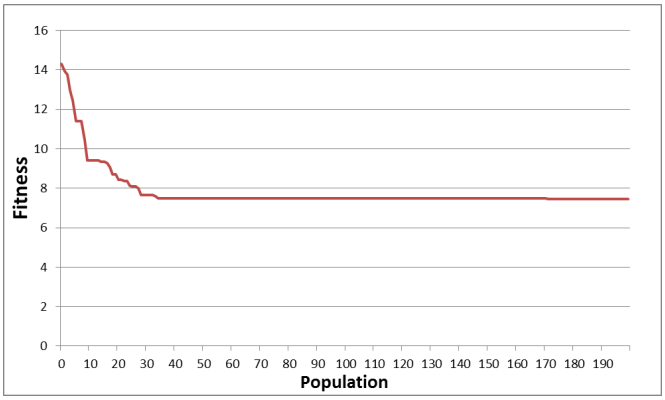

Fig. 5. Solutions improved with every iteration (population)

IV. Conclusion

Our first goal was obtaining a good distribution for public services at developing countries. Technically, the use of genetic algorithms is appropriate, but the major difficulties that we have found to use our algorithms is the inherent lack of basic information at these countries.

In Section II-A shows how the data to work with the genetic algorithm was obtained. This process makes possible to apply the algorithm to countries when the informatios is absent (general case in developing countries). One of the major problems is the rather availability of information about the public transport that connects all isolated places.

Evaluating this solution with the additional common sense based on the knowledge of the studied zone can add some deduced information. Of course, the obtained solution can be improved by adding more related information to our objective function. We can select one solution (the best) among the group of alternatives that we obtain from the algorithm.

According to Figure 5, the results are improving with the iterations of the algorithm. This is notable from the first iteration until around iteration thirty, later the solution keep stable. So, a limit of 30-50 iterations in the algorithm seems to be enough to obtain good results.

In this work, we exposed a basic study case where there is not previous constraints like existing points that must be inserted in the final solution. For future work, more complicated cases can be analized through inserting restrictions in time, distance, applying political criteria, etc.

Acknowledgment

The authors would like to thanks to:

- Universidad Politécnica de Madrid, through its "'XII call for subsidies and grants for cooperative action"' that is financing the project 29_TEDECO_E-GOV-EC. "Identificación de la realidad del Municipio de Santa Elena (Ecuador) para la implantación de e-Government"
- Government of Ecuador through "Secretaria Nacional de Educacion Superior, Ciencia, Tecnologia e Innovacion" (SENESCYT) to support this work.

References

[1] G. Khan, J. Moon, H. Park, B. Swar, and J. Rho, "A socio-technical perspective on e-government issues in developing countries: a scientometrics approach," *Scientometrics*, vol. 87, pp. 267–286, 2011, retrieved: june, 2012. [Online]. Available: http://dx.doi.org/10.1007/s11192-010-0322-5

[2] J. Esteves and R. C. Joseph, "A comprehensive framework for the assessment of egovernment projects," *Government Information Quarterly*, vol. 25, no. 1, pp. 118 – 132, 2008, retrieved: june, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0740624X07000603

[3] H. Carrion, "Internet. calidad y costos en ecuador," 2009, retrieved: june, 2012. [Online]. Available: http://www.imaginar.org/docs/interne_2009.pdf

[4] SUPTEL, "Datos de cuentas y usuarios de internet," December 2011, retrieved: june, 2012. [Online]. Available: http://www.supertel.gob.ec/pdf/estadisticas/acceso_internet_2011.xls

[5] P. S. Bhatnagar, "Egovernment: Lessons from implementation in developing countries," eSocialSciences, Working Papers id:3274, December 2002, retrieved: june, 2012. [Online]. Available: http://ideas.repec.org/p/ess/wpaper/id3274.html

[6] J. H. Holland, *Adaptation in Natural and Artificial Systems*, M. Ann Arbor, Ed. University of Michigan Press, 1975.

[7] D. E. Goldberg, "Genetic algorithms in search, optimization & machine learning," *Addison-Wesley, Reading, MA.*, 1989.

[8] C. Reeves, *Modern heuristic techniques for combinatorial problems*. John Wiley & Sons, Inc., New York, NY, USA., 1993.

[9] R. W. Floyd, "Algorithm 97: Shortest path," *Commun. ACM*, vol. 5, no. 6, pp. 345–, June 1962, retrieved: june, 2012. [Online]. Available: http://doi.acm.org/10.1145/367766.368168

[10] J. T. Alander, "On optimal-population size of genetic algorithms," *Computer Systems and Software Engineering*, vol. 1, pp. 65–70, 1992.

[11] K. A. De Jong, "An analysis of the behavior of a class of genetic adaptive systems." Ph.D. dissertation, Ann Arbor, MI, USA, 1975, aAI7609381.

 ISBN: 978-1-61208-228-8

Understanding the Behavior of the Elderly towards Internet Banking in the UK

Fahad Asmi
Independent Researcher, Pakistan
e-mail: fahadasmi@yahoo.com

Tanko Ishaya
University of Jos, Nigeria
e-mail: ishayat@unijos.edu.ng

***Abstract*—The rapid development and maturity of Electronic Business has stimulated the adoption of Internet Banking in both the developed and developing economies. This adoption is no doubt a result of not just based on the maturity of the Internet, but its ubiquitous nature. As of 2005, about 73% of households have active Internet connection in the UK. At the same time this huge adoption has raised many social and economic issues that should be addressed. This paper presents an investigation into the behavior of elderly people (age segment 55-65) towards Internet banking in the UK. The research constructs were developed based on the 'Decomposed Theory of Planned Behavior' towards providing an understanding of the level of motivations, social influence, perceived usefulness and ease of use for the defined age segment in the UK. Analyzed research output has identified a number of issues that should be addressed in order to increase the involvement of a potential sub-segment of elderly people in e-commerce activities.**

Keywords-Gerontechnology; Digital Divide; Internet Banking and Technology Adoption Model.

I. INTRODUCTION

Technology has changed the whole pattern of life and it is still refining the way we live and do things. Besides its brighter features such as the speed and convenience, Information Technology (IT) has also triggered many complexities requiring research attention. The trend of late adoption of IT related products and services are mostly observed in people who have misconceptions, are techno-phobia and lack IT literacy. The UK is one of the first countries that have adopted electronic mediums for offering government services online. Every institute in UK is using all possible appropriate IT enabled medium of communication for providing their services. Financial institutes are also utilizing different mediums to provide efficient services (such as telephone banking, Automatic Teller Machines (ATM) services and internet banking) for its consumers [34]. This trend is now gaining popularity in most developing economies. However, the opportunities provided by the adoption of the technology also come with a number of new challenges. For example, lack of security [32], loyalty and satisfaction [33]. As a result, a lot of multi-disciplinary research is being undertaken to address these challenges [31].

Existing financial institutes in the UK should address emerging challenges and gaps required for continued improvement of performance and socio-economic growth and development. In the UK, 'healthy life expectancy' is increasing, which creates the need for technological products and services to fulfill the supporting roles for the elder people. Social issues, such as the 'digital divide', are also influencing the society in the UK, where more than 9.2 M are still perceived to be resistant in getting involved in this electronic age. Given the potential use of the Internet in everyday life, there is an urgent need to examine the means to enable older adults to embrace the digital age [9]. Institutes that have plans to formulate their cyber presence should understand the motivation and perception of end-users as key indicators of decision making with respect to product and service development. This paper presents an investigation into the behavior of the elderly (people of an age segment of 55-65) towards Internet banking in the UK.

The remainder of this paper is structured as follows: the next section (Section II), presents an analysis of the related background information required for a justification for the research presented in this paper. It describes the development of the model used. Section III presents a description of the research objectives and the main methodology used and the analysis of the findings presented in Section IV. Section V concludes with ideas for future work.

II. BACKGROUD

The background to this paper is based on three related areas of work - including an evaluation of ICT development in the UK, its potential benefits and issues affecting its full adoption by the elderly in Section IIA. It also presents a generic review of ecommerce in the UK in Section IIB and a study of the behavior of the elderly consumers to ecommerce in Section IIC. Finally, it reviews a number of innovation adoption models that helped in defining a theoretical framework for this research in Section IID.

A. IT/ICT and Gerontechnology in UK

History reveals how Germany and Japan took over leadership in manufacturing industry, which was once UK's identity [1]. Presently, India and China seem to be grabbing it from Germany and Japan [1]. Thus, to re-develop a resilient UK economy, intensive investment in IT and ICT sector has been made. Its objective is to make UK a 'virtual hub' for the globe [1]. The Internet is developing a uniformed global culture and the UK is greatly participating in it, demonstrated by the 19 million out of 25 million people that have Internet connections at home in UK [19]. As a result, Internet economy is the 5th largest part of the whole UK's economy [19].

 ISBN: 978-1-61208-228-8

ICT has opened new ways for elderly people to perform roles and responsibilities at home and work in progressive society as active and independent citizens with the support of developed ICT innovations [38]. For example, improvements in health care procedures while reducing its cost [37]. The adoption of ICT also encourages individuals to maintain their privacy and security [37]. With the increase in age, people experience different types of illnesses - including visual problems, hearing impairments and arthritis [40], and it is clear that ICT innovation can help these people to perform self-care supporting activities. Researchers have realized the heterogeneity of grey market in terms of their age group, interests, skills and abilities that can be useful in the phase of learning and adopting ICT innovation in older people. [39].

The demand of ergonomics and assistive technology for elderly people in 1990's presented research and development opportunities and challenges required to fulfill future needs [41]. It is the output of multi-disciplinary overlapping research, where 'Technology' interacts with 'Gerontology' to work for betterment of the life of elderly people [43]. For example, aging issues decrease the chances for elderly people to continue their routine activities at the same pace. The technology that can help to avoid or decrease such chances for elder people is classified as 'gerontechnology' [14]. Gerontechnology also encourages the involvement of elderly people during the design and development phase of any product and services for their use [42]. Government and educational institutes in UK are contributing in all possible ways to cope with uprising aging issues of society by re-engineering and re-designing operations and management.

B. E-commerce in UK

Electronic use in all business processes has re-shaped all sets of actions and enhanced the business growth to higher levels [18]. Financial firms are using latest technological support to increase transactional security and provide maximum number of mediums for serving consumers [4]. In the framework of e-commerce, Internet banking is defined as a major part of its applications [18], which still has many acceptance issues in consumer's mind [7]. Internet banking can be classified broadly in 3 different forms: (1) To only provide information about services; (2) To be used as communication channel only but can't access bank's internal network; (3) Used to offer complete operational features and services through electronic platform [22]. In electronic banking, first remarkable development is the ATM, which extraordinarily reduced the cost of every transaction [5]. At the same time, the hesitation to use ATM by older people has created high level of attention [29]. The Bank of Scotland is the first financial institute in UK that started providing electronically in 1980's [10]. By the year 2006, about 17 million adult population of the UK adopted Internet banking, which is about one thirds (1/3) of the total population of adults in UK [6]. The rate of adoption in the case of Internet banking can be faster and higher if the crucial factors like privacy, legislation and security are addressed [21].

C. Elderly People in UK

Studying consumer behavior of specific segment of people provides us with useful information about them. It can be used to predict the products and services that they need [9] i.e.; older people usually buy products and services according to their convenience [28]. In the UK, the well off people usually starts to have leisure activities in their older lives [13], and they are concerned more about health and mobility related issues [8]. Due to heterogeneity in terms of attitude and values of mature market, it is important to subdivide them, usually researchers divide them as 55-64, 65-74, 75-85 and older than 85 [13]. In the EU, people of age 45 and above are increasing as compare to the adults of age 45 or less [16]. Still, in this era of technological revolution, people prefer to have face-to-face communication where possible; this behavior can be understandable [20], given that elderly people have phobia of using technological products and services [17].

D. Theoretical Frameworks for Technology Adoption

To understand the behavior of elderly people (age segment 55-65) towards Internet banking in the UK, a technology adoption model is required which can figure out the consumers' perceptions and attitude. As a quick flashback of existing models, 'Innovation Diffusion Theory' (IDT) can be considered as a first proposed model that defined the procedure of accepting or rejecting any new innovation [26]. Afterwards, many social psychological researchers suggested many variables to predict behavior i.e. attitude, social norms, subjective norms and perceived behavioral control. Specifically for technology adoption 'technology acceptance model' (TAM) can be considered as the first model to understand technology adoption in consumers. TAM defines 'perceived ease of use (PEOU) and 'perceived usefulness' (PU) as major predictors of any technology adoption [11]. Similarly, the 'decomposed theory of planned behavior' (DTPB) was developed by modifying the 'theory of planned behavior' (TPB) to understand technology adoption behavior [30]. DTPB describes (1) 'Attitude' as sum of compatibility, perceived ease of use and usefulness, (2) 'Subjective norms' as a composition of peer and superior influences and (3) 'Perceived behavioral control (PBC)' as the output of self-efficacy, resource and technology facilitating conditions [2]. Researchers have already used different variables to make an extension of existing DTPB for research i.e., 'risk' [15] and 'friendliness' [25] as a part of DTPB model.

Based on the analysis of existing models, Figure 1 presents a theoretical model adopted for this research.

 ISBN: 978-1-61208-228-8

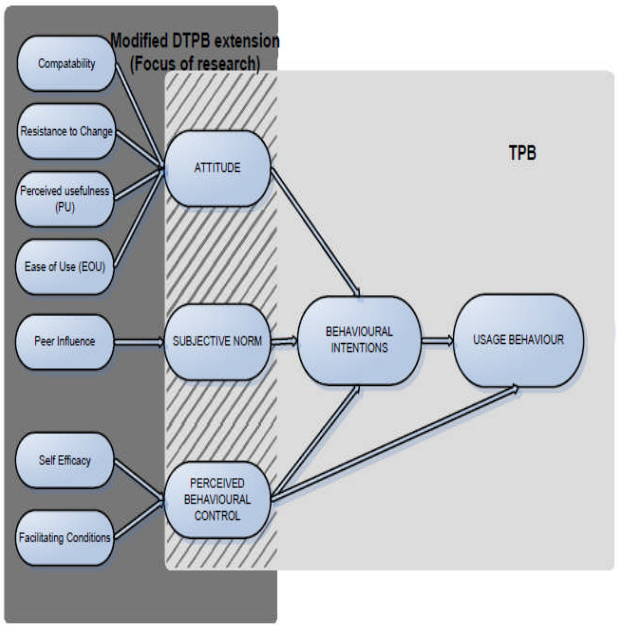

Figure 1: Proposed model for the current research.

In the current research, DTPB model has been adopted and 'resistance to change' added as a part of the variables required for predicting 'attitude' as it has been done in many previous researches [35, 36]. 'Superior influence' will be eliminated from the variable to define 'subjective norms' and only peer influence will be discussed. PBC will be analyzed on the basis of available facilitating conditions and individual's self-efficacy to perform IB related tasks and operations in the current research model. More specifically, only the relationship among variables will be analyzed which are defining 'attitude', 'subjective norms' and PBC of elderly people (age group: 55-65) towards Internet banking in the UK and will not explain the relation of 'intentions' and 'behavior' with 'attitude', 'social nom' and PBC

III. OBJECTIVE OF THE CURRENT RESEARCH

The purpose of the current research is to extract factors that have led to the low adoption Internet banking by the elderly people (with age segment of 55-65) in the UK, in the following dimensions:

- Cross-interaction among 'resistance of change', PEOU, PU and 'Compatibility'.
- Impact of 'social influence' in the case of IB.
- How 'self-efficacy' and 'facilitating conditions' affect the PBC of elder people (age segment: 55-65) towards Internet banking.

A. Methedology

For achieving the goal, philosophical stance has been 'Interpretive' due to the focus on the complex social issues. The research design is 'exploratory' since the aim is to describe and discuss problem that potentially leads to an increase in knowledge. For conducting research, quantitative-qualitative methods have been adopted, by conducting personal, semi structured, face-to-face interviews and surveys in the form of questionnaires. More specifically, a sample size of a population 100 users (70 non-users and 30 uses of Internet Banking) has been collected according to the simplified random sampling method. On the other hand, convenience sampling was used for conducting interviews with 8 participants (5 non-users and 3 users). Inductive-deductive study approach helps to work under a framework and suggests new elements as a part of existing model. Qualitative data was gathered and analyzed by making transcripts of interviews, categorizing according to the variables and identifying relations in the summarized data. The captured quantitative data was analyzed using descriptive statistics to define and challenge existing relationships among variables.

B. Data

The table below presents the data collected from responses of both users and non-users of Internet banking.

TABLE 1: DOMINATING FEATURES COLLECTED AS RESPONSE DURING THE SURVEY FROM USERS AND NON-USERS OF IB.

Questions	**Response types**	**Non-users of IB**	**Users of IB**
IB can make my banking activities easier?	Agree	20%	58%
	Neutral	35%	26%
	Disagree	45%	16%
Banks are providing sufficient information to adopt IB?	Agree	19%	38%
	Neutral	27%	36%
	Disagree	55%	26%
E-transactions are secure?	Agree	27%	42%
	Neutral	38%	38%
	Disagree	35%	20%
Have time to get involve in IB?	Agree	7%	56%
	Neutral	39%	36%
	Disagree	54%	8%
Discuss finance related issues with family and close friends?	Agree	22%	6%
	Neutral	24%	26%
	Disagree	74%	68%
People perceived IB includes complex transactional procedures?	Agree	50%	40%
	Neutral	32%	26%
	Disagree	18%	34%
IB provides personalized facility	Agree	15%	40%
	Neutral	35%	25%
	Disagree	50%	35%

 ISBN: 978-1-61208-228-8

IV. FINDINGS AND ANALYSIS

Demographic profile of IB non-users in the sample is shown in Table 2.

TABLE 2: DOMINANT FEATURES OF COLLECTED NON-IB USERS SAMPLE.

Age Segment	55-60	45%	13% have not specified any option
	60-65	42%	
Gender	Male	48%	4% have not specified any option
	Female	48%	
Occupation	Employed	65%	6% holds other options
	Retired	28%	
Educated	Secondary	48%	Remaining 37% holds other option
	Primary	15%	
Respondents that used computers before	82% of non-users of IB used computers before		
Respondents that used the Internet	More than 65% used Internet		

Major questions used to define each variable are following:

Compatibility
- I interact with technology in routine life.
- IB adoption is possible for me.
- IB can improve my banking activities.

Resistance to Change
- I always look for new developments to make my life easy.
- I always look for new technology or services.
- Banks should adopt new technologies.

Perceived Usefulness
- IB can make my life comfortable.
- IB is valueable service.
- IB is personalized service.
- E-transactions are secure mode of payments.

Perceived Ease of Use
- Learning IB is an easy process
- IB's transaction handling is risky procedure.
- Online transaction procedure is easy.

Facilitating Facilities
- I have computer and Internet facility.
- I am frequent Internet user.
- I have time to involve in IB.
- I have knowledge of IB.

Self-efficacy
- I have confidence to search over the Web.
- I can understand the e-transaction procedures easily.
- I can learn technological interaction easily.

Peer Influence:
- I usually try things what my friends and family suggests.
- Discuss about plans to buy with my family and friends.
- Updates about new technology or services make me happy.

After complete reliability and validity test, logical division of the current area of the paper has focused on identifying aspects of attitude, PBC and social norms of elder people towards Internet banking. A blend of quantitative and qualitative data was used to support the arguments.

Reliability and validity tests are important for verifying relationships between interrelated items. Value of Cronbach Alpha is used for this purpose, which should be higher than .70. If the items are less than 10, inter item correlation can be measured which should be ranged between .2 - .4 [23], as shown in the Table 3.

TABLE 3. RELIABILITY AND VALIDITY TEST

Area of Study	Number of Items (Questions involved)	Cornbach's Alpha	Inter-item correlation (mean)
Compatibility	6	0.761	0.346
Ease of Use	3	0.592	0.340
Perceived Usefulness	6	0.605	0.207
Resistance to Change	5	0.738	0.300
Facilitating Facilities	5	0.713	0.339
Self-efficacy	6	0.730	0.330
Peer and Family Influence	4	0.511	0.303

A. Attitude

The positive correlation is observed among 'Compatibility', PU and PEOU. Whereas negative association of 'resistance to change' is observed with all other 'Attitude' related variables in the case of elder people towards Internet banking, as shown in Figure 2.

 ISBN: 978-1-61208-228-8

Correlations

		useful	easeofuse	compatability	resistance
useful	Pearson Correlation	1	.707**	.432*	-.787**
	Sig. (2-tailed)		.000	.035	.000
	N	35	35	24	35
easeofuse	Pearson Correlation	.707**	1	.757**	-.553**
	Sig. (2-tailed)	.000		.000	.001
	N	35	35	24	35
compatability	Pearson Correlation	.432*	.757**	1	-.346
	Sig. (2-tailed)	.035	.000		.098
	N	24	24	24	24
resistance	Pearson Correlation	-.787**	-.553**	-.346	1
	Sig. (2-tailed)	.000	.001	.098	
	N	35	35	24	35

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 2: Correlation analysis of 'Perceived Ease of Use', 'Perceived usefulness, 'Compatibility' and 'Resistance to Change' in the case of IB for target segment.

1) Compatibility

Compatibility is the degree at which any improvement can stick with existing set of values and needs. Initial quantitative findings from the non-users of IB, related to compatibility are mentioned in Table 4.

TABLE 4: 'COMPATIBILITY' RELATED FEATURES OF NON-USERS OF IB

Having computer literacy	50%
Having internet literacy	45%
Intentionally avoids IB	75%
Having ability to perform tasks of IB	60%

The trend of adopting Internet banking in male's increases with age but this trend is converse in females.

During qualitative analysis, questions asked about involvement of technology in financial activities, responders replied with the least usage of ATM for withdrawing money but more interestingly, people still use Internet for web browsing and emails. Descriptive statistics shows that there is a positive relation of 'usefulness' and 'ease of use' with 'compatibility' in the current case of study. Actions from socially active firms to decrease digital divide, increase in retirement age and higher ratio of males as compare to females in working environment, can be the reason of output of this research. Literature and current research defines that, people having the willingness, better education, need of IT for job roles and it in a values, are adopting better way.

2) Perceived Ease of Use (PEOU)

PEOU defines how easy it is to adopt a technology. Quantitative research have presented a number of observed patterns: (1) 34% of people assume IB is difficult to adopt and 40% have neutral behavior towards IB. (2) 45% of non-users defines IB procedures as complex suite of actions and 40% have never thought about IB. (3) Mostly employed people have neutral behavior about IB and retired people have negative perception of it. Statistical analysis summarized 'PEOU' as having a positive relation with PU and compatibility. Qualitative analysis defined that target segment has knowledge about IB but they usually felt more comfortable in face-to-face communication. Literature supports the findings on the basis of following reasons. (1) Existing measures of quality of e-service highlight the importance of 'enjoyment', 'responsiveness' and 'contact' [24]. (2) A fewer packages of information provided by financial institutes about IB in target segment can be the reason of 'neutral' and negative' behavior towards IB. (3) Existing techno-phobia, myths and a huge communication gap can be the reason of this survey's response.

3) Perceived Usefulness

This explains how any innovation can improve and simplify the tasks. Quantitative survey indicates (1) 35% of non-users have 'neutral' perception about IB. (2) More than 85% of non-users have privacy issues. (3) 50% of non-users defined complexities in online transactional process as a reason to avoid IB. (4) 35% have doubts about IB security. Still more than 37% people have 'neutral' behavior towards IB. Descriptive statistics concludes positive relation of PU with PEOU and Compatibility. While conducting interview with an IB user, he mentioned the security checks provided by his bank for each transaction as a reason of his satisfaction. Similarly, one of the non-users of IB indicated security as one of the main factors that made them to avoid IB. All observations concluded a few elements, which should be address. (1) Implementations of IB at information, communication and transactional level should be re-defined to remove 'lack of trust' and 'critical security issues' to attract consumers [22]. (2) In the case of IB, there can be a chance that people experienced value is less as compare to the perceived value which made them non-users [24].

4) Resistance to Change

Survey concludes that 50% people avoid IB because of complex set of actions for each transaction. The following 2 patterns support the argument that the target segment are potential IB consumers, (1) 55% of non-users are willing to adopt new innovation to improve life support. (2) 28% of non-users are not avoiding IB because of complex transactional process. Descriptive statistical analysis concludes that PEOU and PU have positive relation with resistance of change, but the current research hasn't found any relationship between 'resistance to change' and 'compatibility'. One interviewer indicated that they have not used the Internet for IB purpose and another interviewer rejected IB to adopt at any cost. Literature defined resistance towards IB because of following reasons, (1) Non-users in target segment avoids IB because they di not take risks and challenges, do things only on the basis of 'loss based selection' improve life in restricted dimensions [12]. (2) Everybody avoids testing any new idea where the chances of financial loss are high. To conclude the set of variables that are defining 'Attitude' are Inter-related and have impact on each other.

B. Peer influence (Social norms)

In the current research, only 'peer-influence' will be discussed on how friends and family effects any financial

 ISBN: 978-1-61208-228-8

activity in target segment, statistical analysis can be summarized with the following dominant features. (1) For general shopping, 40% usually attempt the recommended products and services. (2) 60% of the sample hasn't ever been advised to use IB. (3) Educated people have interest in recent technological developments. On the other hand, qualitative survey also showed the same patterns, i.e., an IB user recommended IB to other friends and family and they find it easy and convenient to adopt. Perspective of non-users is different, interviews concludes that people usually avoid discussing financial concerns with friends and family. From literature perspective the following have been observed:

- IB provides less opportunity for trial-ability and observe-ability [26].
- People with 'neutral' opinion can be attracted by proper encouragement and support.
- People only pays attention to the information, which supports their opinion, such biased attitude can be the reason of avoiding IB.

C. *Perceived Behavioral Control*

Research have shown a possitive relationships between all the factors. For example, in the third compnent, the relationship between 'facilitating facilities' and 'self-efficacy' to define PBC towards IB for the target age segment was invetigated using the the Pearson Product-momment correlation coefficient (*r*). From the results in Figure 3, the result *r*=.760 indicates a strong, positive correlation between the two variables.

Correlations

		facilities	self_efficacy
facilities	Pearson Correlation	1	.760**
	Sig. (2-tailed)		.000
	N	24	24
self_efficacy	Pearson Correlation	.760**	1
	Sig. (2-tailed)	.000	
	N	24	24

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 3: Correlation analysis of 'Facilitating facilities' and 'Self efficacy'.

1) *Facilitating Facilities*

It includes the know-how about operation in IB and physical devices, which are needed for IB. From questionnaires, following attributes can be concluded, (1) 60% of non-users have knowledge about IB. (2) Employed and educated people have more knowledge about IB compared to the rest of the sample. Descriptive analysis concludes facilitating facilities have positive relation with self-efficacy. On the other side, qualitative analysis says that more interaction of non-users with technology can reduce myths and techno-phobia. Findings are understandable in the following ways. (1) IB has less 'Tangibility' factor. (2) Learning ability of IB decreases with age [3]. (3) Intensive attention should be paid on each level of IB (Transactional, communication and information).

2) *Self-efficacy*

This is about an individual's ability to do tasks with confidence. In the current target segment, self-efficacy is getting low with increase in age. Similarly, their 'neutral' behavior towards IB is also getting 'negative' with increase in age. As it is previously discussed that presence of facilitating conditions can increase the self-efficacy in target segment towards IB. Quantitative and qualitative survey concludes that PBC is not the major reason to avoid IB. Target segment have self-efficacy and facilitating conditions. Literature defines that elder people only adopt new innovation when they have any problem in existing pattern of work and target segment is quite satisfied with traditional face-to-face banking [27]. Secondly, Most of the people have the ability to do different and unique tasks, they only lacked the proper guidance and instructions to adopt them [3], and this can be the case for the target age segment.

V. CONCLUSION AND FUTURE RESEARCH

During the whole span of quantitative and qualitative research, target segment prefers traditional banking system where they have face-to-face interaction. While studying behavior of elderly people towards Internet banking 'compatibility' is observed the most influencing as compare to PEOU and PU to form attitude. 'Resistance of change' has negative relation with all variables who are composing 'attitude'. In the behavior towards IB, Social norms in the form of 'peer influence' shows less influencing variable as compare to the impact of 'attitude' in defining behavior towards IB in the elderly people in the UK. In the collected quantitative sample, more than 45% of non-users of IB in target segment have most of the facilitating conditions and have enough self-efficacy to perform IB related tasks. The age segment 55-65 with education and employment should be address by financial institutions to increase their motivation, and encourage them to participate in IB services. The ability to learn and observe exists in elder people in the UK but financial institutes haven't ever interacted with them while taking care of their heterogeneity.

The current research used a descriptive statistical approach for analyzing the relationship among variables to define attitude, social norms and PBC. Multiple-regression can be applied to make the research more authentic. In target segment, definitions of 'perceived risks', 'quality' and 'privacy' are different in the case of IB, so it is needed to be re-defined for target segment's needs. A a new scale for measuring the quality in the case of IB can be helpful to predict better adoption rate. Longitudinal in terms of time horizon can enrich the findings of the current research as technology adoption decisions includes behavioral and psychological dimensions

REFERENCES

[1] J. Acklam and C.Chang, "ICT-enabled UK-based Global Virtual Hubs," BCS Management Forum Strategic Panel, 2009, Available

 ISBN: 978-1-61208-228-8

online: www.bsc.org/upload/pdf/global-virtual-hubs-paper.pdf BCS [retrieved: January, 2011]

[2] K.I. Al-Qeisi, "Analyzing the Use of UTAUT Model in Explaining an Online Behaviour: Internet Banking Adoption," PhD Thesis, 2009, Brunel University.

[3] A. Bandura, "SELF EFFICACY: the exercise of control," New York: W.H. Freeman and Company, 1997

[4] E. Bock, "The introduction of e-Commerce systems in the banking industry - an example of the BBBank and some empirical results," Proc. ECIS, 2000

[5] F. Calisir and C.A.Gumussoy, "Internet banking versus other banking channels: Young consumers' view," International Journal of Information Management. 2008, vol.28(3), pp. 215-221.

[6] H. Celik, "What determines Turkish customers' acceptance of internet banking?," International Journal of Bank Marketing. Vol.26(5), 2008, pp. 353-370.

[7] Y.T. Cheng "Dynamics of Internet Banking Adoption," Available online: -- [retrieved: May 2011]

[8] I. Chaston, "Boomer Marketing: Selling to a Recession Resistant Market," London: Routledge, 2009.

[9] B. Clegg, "The invisible customer: strategies for sucessive customer service down the wire. London: Kogan Page, 2000.

[10] E. Daniel, "Provision of electronic banking in the UK and the Republic of Ireland," International Journal of Bank Marketing. vol.17(2), 1999, pp.72-82.

[11] F.D.Davies "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," MIS Quarterly, vol.13 (3), 1989, pp. 319-340.

[12] P.S. Fry and C.L.M. Keyes, "New Frontiers in Resilient Aging: Life-strengths and well-being in Late Life," Cambridge: Cambridge University Press, 2010.

[13] B. Gunter, "Understanding the older consumer: the grey market," London: Routledge, 1998

[14] T.L. Harrington and M.K. Harrington "Gerontechnology: Why and How," Eindhoven, Netherlands, Shaker Publishing B.V, 2000

[15] E. Huang and M.H.Chuang "Extending the theory of planned behaviour as a model to explain post-merger employee behaviour of IS use," Computers in Human Behavior, vol.23, 2007, pp. 240-257.

[16] D. Jobber, "Principles and Practice of Marketing," McGraw Hill Higher Education, 2006

[17] N.Joseph, and B. Stone "Diagnosis: Interent Phobia [Seniors who shun the web miss out. How to help]," unpublished .

[18] R. Kalakota and A.B. Whinston "Electronic commerce: a manager's guide," 1st ed, Addison-Wesley, 1997

[19] C. Kalapesi, et al, "The connected Kingdom: How the internet is transforming the U.K. economy," Boston, MA: The Boston Consulting Group, 2010

[20] M.Lindstrom and T.F Andersen, "Brand building on the Internet", Cambridge: Kogan Page, 2000

[21] A. Y. Mermod, "Customer's Perspectives and Risk Issues on E-Banking in Turkey; Should We Still be Online?," Journal of Internet Banking and Commerce, vol. 16(1), 2010, pp-1-15

[22] OCC, "Internet Banking: Comptroller's Handbook. Office of the Comptroller of the Currency," unpubshed

[23] J. Pallant, SPSS: Survival Guide, 2007

[24] A. Parasuraman, V.A. Zeithaml, and A. Malhotra, "E-S-QUAL A Multiple-Item Scale for Assessing Electronic Service Quality," Journal of Service Research. Vol.7(10), 2005, pp. 1-21.

[25] P.E. Pedersen, "Adoption of Mobile Internet Services:An Exploratory Study of Mobile Commerce Early Adopters," Journals of Organization computing and electronic commerce, Vol.15(2), 2005, pp. 203-222.

[26] E.M. Rogers, "Diffusion of Innovations," New York: Simon and Schuster, 2003

[27] R. Sandra-Cornett, "Teaching the Elderly," Ohio: The Ohio State University Medical Center, 2001

[28] J.N. Sheth, R.S. Sisodia, et al., "Impact of Demographic Diversity and Information Technology on the Future of Consumer Behavior," Journal of the Academy of Marketing Sciences. MSI/JAMS Special Issue, 1999, pp. 1-28

[29] A. Tarakanov, "Use and non-use of Automatic Teller Machines by older people in Israel," Journal of the International Society for Gerontechnology, vol.3(2), 2004, pp. 107-110.

[30] S. Taylor and P.A.Todd, "Understanding information technology usage: a test of cometing models," Institute of Operations Research and Management Sciences, vol.**6**(2), 1995, pp. 144-176.

[31] M.A. Sarlak and A.A. Hastiani, "E-Banking and Emerging Multidisciplinary Processes: Social, Economical and Organizational Models," Business Science Reference, Hershey, New York, 2010

[32] S. Drimer, S.J. Murdoch, and R. Anderson, "Optimised to Fail: Card Readers for Online Banking," Financial Cryptography and Data Security. Springer-Verlag, 2009, pp.184-200.

[33] A. Floh and H. Treiblmaier, "What Keeps The E-Banking Customer Loyal? A Multigroup Analysis Of The Moderating Role Of Consumer Characteristics On E-Loyalty: In The Financial Service Industry," Journal of Electronic Commerce Research, Vol. 7 (2), 2006

[34] ICT and e-Business Impact in the Banking Industry A Sectoral e-Business Watch: European Commission, Enterprise & Industry, 2008: [Available online: http://ec.europa.eu/enterprise/archives/e-business-watch/studies/sectors/banking/documents/Study_06-2008_Banking.pdf]

[35] V.T. Ragoobur and A. Ayrga, "Is Mauritius Ready to E-bank? From A Customer and Banking Perspective," Journal of Internet Banking and Commerce, vol. 16(1), 2011

[36] M. Sathye, "Adoption of Internet Banking by Australian consumers: An Emperical investigation," International Journal of Bank Marketing, vol. 17(7), 1999, pp. 324-334.

[37] G. Kathrin and C. Michael, "ICT enabled independent living for elderly : A status-quo analysis on products and the research landscape in the field of Ambient Assisted Living (AAL) in EU-27," Institute for Innovation and Technology, 2010.

[38] F. Amaro and G. Henrique "ICT for Elderly People: «Yes, 'They' Can!»," Proc. International Conference on e-Commerce, e-Administration, e-Society, e-Education, and e-Technology, Tokyo, Japan, 2011, pp-3792-3803 URL: http://hdl.handle.net/10400.11/820

[39] S. Sayago, P. Fobers, and J. Balt, "Old people becoming successful ICT learners over time: Challenges andstrategies through an ethnographic lens," Educational Gerontechology, in press

[40] Chronic Conditions A challenge for the 21st century: Chronic And Disabling Conditions Challenges For The 21st . Century. Washington, DC, National Academy On An Aging Society.

[41] J.E.M.H. van Bronswijk, H. Bouma, and J.L. Fozard, "Technology for quality of life: an enriched taxonomy," Gerontechnology, vol.2 (2), 2002, pp-169-172

[42] J.A.M. Graafmans, J.L. Fozard, J. Rietsema, A. van Berlo, and H Bouma,"Gerontechnology: Matching the technological environment to the needs and capacities of the elderly," In Brookhuis, K. A. (Ed. The European Chapter of the HFES Tenth Anniversary Meeting. Soesterberg, 1996, pp-19-31

[43] B. Dara-Abrams, "Toward a Model for Collaborative Gerontechnology: Connecting Elders and their Caregivers," Proc. Sixth International Conference on Creating, Connecting and Collaborating through Computing, 2008, pp-109-114 : DOI: 10.1109/C5.2008

 ISBN: 978-1-61208-228-8