Sérgio Tenreiro de Magalhães
Hamid Jahankhani
Ali G. Hessami (Eds.)

Communications in Computer and Information Science 92

Global Security, Safety, and Sustainability

6th International Conference, ICGS3 2010
Braga, Portugal, September 2010
Proceedings

Springer

Communications
in Computer and Information Science 92

Sérgio Tenreiro de Magalhães
Hamid Jahankhani Ali G. Hessami (Eds.)

Global Security, Safety, and Sustainability

6th International Conference, ICGS3 2010
Braga, Portugal, September 1-3, 2010
Proceedings

Springer

Volume Editors

Sérgio Tenreiro de Magalhães
Universidade Católica Portuguesa
Braga, Portugal
E-mail: stmagalhaes@braga.ucp.pt

Hamid Jahankhani
University of East London
London, UK
E-mail: hamid.jahankhani@uel.ac.uk

Ali G. Hessami
City University
London, UK
E-mail: alihessami@aol.com

Library of Congress Control Number: Applied for

CR Subject Classification (1998): C.2, K.6.5, D.4.6, E.3, H.4, J.1

ISSN 1865-0929
ISBN-10 3-642-15716-5 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-15716-5 Springer Berlin Heidelberg New York

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper 06/3180 5 4 3 2 1 0

Preface

The annual International Conference on Global Security, Safety and Sustainability (ICGS3) is an established platform in which security, safety and sustainability issues can be examined from several global perspectives through dialogue between academics, students, government representatives, chief executives, security professionals, and research scientists from the United Kingdom and from around the globe.

The three-day conference focused on the challenges of complexity, rapid pace of change and risk/opportunity issues associated with modern products, systems, special events and infrastructures. The importance of adopting systematic and systemic approaches to the assurance of these systems was emphasized within a special stream focused on strategic frameworks, architectures and human factors. The conference provided an opportunity for systems scientists, assurance researchers, owners, operators and maintainers of large, complex and advanced systems and infrastructures to update their knowledge on the state of best practice in these challenging domains while networking with the leading researchers and solution providers.

ICGS3 2010 received paper submissions from more than 17 different countries in all continents. Only 31 papers were selected and were presented as full papers. The program also included a number of keynote lectures by leading researchers, security professionals and government representatives.

June 2010 Hamid Jahankhani

Organization

Organizers

Universidade Católica Portuguesa, Faculdade de Ciências Sociais, Portugal
University of East London, School of Computing, IT and Engineering, UK
IEEE SMC, Technical Committee for Systems Safety & Security

General Chair

Hamid Jahankhani	University of East London, UK

IEEE SMC Chair

Ali Hessami	Vega Systems, UK

Program Committee Chair

Sérgio Tenreiro de Magalhães	Universidade Católica Portuguesa, Portugal
Henrique M.D. Santos	University of Minho, Portugal

Local Chair

Sérgio Tenreiro de Magalhães	Universidade Católica Portuguesa, Portugal

Program Committee

Major G.S. Aki Huhtinen	National Defence College, Finland
Ali Hessami	Vega Systems, UK
Ali Sanayei	Director of ITM Research group, University of Isfahan, Iran
Ameer Al-Nemrat	University of East London, UK
Carlos Sousa Pinto	University of Minho, Portugal
David Lilburn Watson	Forensic Computing Ltd., UK
Duarte Duque	EXVA Ltd., Portugal
Elias Pimenidis	University of East London, UK
Feng Hsu	Aerospace Technology Working Group, USA
Col Fernando Freire	Military Academy, Portugal
Florin Gorunescu	University of Medicine and Pharmacy of Craiova, Romania
Filipe de Sá-Soares	University of Minho, Portugal

George Weir	University of Strathclyde, UK
Gianluigi Me	University of Rome Tor Vergata, Italy
Henrique M.D. Santos	University of Minho, Portugal
Hossein Jahankhani	University of East London, UK
Hugo Gamboa	New University of Lisbon, Portugal
Kenneth Revett	University of Westminster, UK
Konstantinos Kardaras	Technical Consultant, Greece
Leonel Duarte dos Santos	University of Minho, Portugal
Marios Efthymiopoulos	University of Cyprus, Cyprus
Orhan Gemikonakl	Middlesex University, UK
Mohammad Dastbaz	University of East London, UK
Paulo Viegas Nunes	Military Academy, Portugal
Rain Ottis	Cooperative Cyber Defence Centre of Excellence, Estonia
Reza Sahandi	Bournemouth University, UK
Sérgio Tenreiro de Magalhães	Catholic University of Portugal, Portugal
Sufian Yousef	Anglia Ruskin University, UK
Vasilios Zorkadis	Directorate of the Hellenic Data Protection Authority, Greece

Table of Contents

Management of the Benefits on the Client's Involvement on Ergonomic Analysis

Isabel F. Loureiro*, Celina P. Leão, and Pedro Arezes

Departamento de Produção e Sistemas, Escola de Engenharia, Universidade do Minho, 4710-057 Braga, Portugal

Abstract. Nowadays, market trade economy is witnessing to a continuous development and transformation. The organizations come to be seen as socio-technical systems with new ergonomic contexts. Various types of relationships can be established. From the ergonomic analysis point of view, it is necessary to understand all the mechanisms that regulate these relationships. The interaction between clients and professionals (workers) reproduce a relationship that can be important to the ergonomic analysis. This paper allows a better comprehension of the relationship in the effective's ergonomic intervention. A case study was analyzed in a private health sector using the Ergonomic Three-dimension Analysis as an ergonomic approach. This analysis is made by three different but related dimensions: analyst, professional and client. The results show that that clients' involvement in the ergonomic analysis will benefit the ergonomic intervention and consequently the professional environment.

Keywords: systems, relationships, clients, analysis, intervention.

1 Introduction

In the last years, the corporations are no longer the centre of the market economy. They became a scene for different actors who, over time, have different roles within organizations. This new market approach provides that organizations come to be seen as a socio-technical system [1] comprising a set of different but interrelated subsystems involved: supplier, customer, employee, patient, managers … [2]. The oncoming between these subsystems is the result of the technological revolution boom involving new communication technologies and information [3]. This whole process of changing will have impact in the chain distribution were clients are assuming a vital role.

According to Lindon (2000) [4], the corporation is constructed in such a way as to ensure an effective response to the marketing exigencies. Client commands the corporation destiny, and the top management strategies are developed in horizontal collaboration, stretching the different organization hierarchies [4]. The quality of goods and services are improved through the participation of individuals at all organization levels [5].Therefore, the total quality management philosophy must be focused not only in workforce satisfaction, but also in clients' satisfaction since in modern social-technical systems they are intrinsically linked to the organizations. This complexity of work systems [6] originates new ergonomic contexts, and for each one, it can be establish various types of relationships:

* Corresponding author; ID2500@alunos.uminho.pt

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 1–8, 2010.

workers relationships, supervisors/workers relationships, administrative relationships with users, with health care patients and psycho-social education, and clients' relationships. The clients business relationships that results from concessions between the corporation, the worker, the environment and the promising client/consumer, are important in the development on the present work [7]. From the ergonomic analysis point of view, it is necessary to understand all the mechanisms that regulate these relationships. The recognition that the customer, client or user is an active part of the new ergonomic contexts simultaneously with the emergence of new Commercial Areas with Free Circulation of People (CAFCP) requires the development of new ergonomic approaches. These, should allow a more detailed analysis of the real activities of individuals, considering the CAFCP not only in an occupational perspective, but also with a usability point of view. Effectively, the CAFCP are common areas where clients and professionals interrelate and can equally be exposed to the same ergonomic risk.

In this paper, it is proposed an evolution for the traditional ergonomic occupational analysis where the binomial constituted by the dimension of the Ergonomist (analyst) and employee gives rise to the trinomial composed by the dimensions of the professional, analyst and clients (Figure 1) [8].

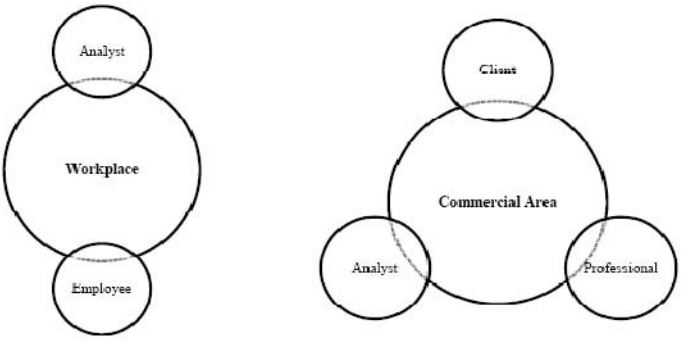


Fig. 1. Binomial *(left) versus* Trinomial *(right)*

The ETdA model is a new methodology that allows the ergonomic analysis and intervention in CAFCP. This is a three-dimensional analysis made by three different but related dimensions: analyst, professional and client. The interaction between clients and professionals (workers) reproduce a relationship that can be important to the ergonomic analysis. On one hand, it is considered that the answers given by professionals in the evaluation forms may not match the vision of real work, since the Human Being is a bad estimator of itself, therefore the professional evaluation may be overestimated or underestimated by him [8]. This paper proposes that clients' results may serve to correct these deviations contributing to the business strategies improvements. On the other hand, it is assumed that clients' dimensions results make easier the ergonomic intervention since the oncoming of this dimension to the management creates co-responsibility in the changes to be implemented.

A case study was analysed in a private health sector using the ETdA model as an ergonomic approach [8]. The commercial areas with free circulation of people where correctly identified and the three dimension observation tolls developed [9] allow a

better comprehension of the relationship Client/professional in the effectives of the ergonomic analysis and intervention.

2 Methodology

This study was carried out among the health sector, in six parapharmacies located in different geographic regions in the North Portugal. Several factors had contributed to the six parapharmacies selection, namely location, availability, contact and accessibility [8]. Table 1 describes the observation tools used in each ETdA dimension [9].

Table 1. Observation tools used in ETdA dimensions

Dimension	Observation tool
Client	Questionnaire
Professional	Evaluation forms
Analyst	Direct and indirect observation

The evaluation forms and the analyst direct and indirect observations used for the professional and analyst dimension, respectively, are the observation tools used in the ETdA model. They follow the methodology used in the Ergonomic Workplace Analysis- EWA [10], and were correctly adapted to be applied in this study-case [11]. This methodology allows a systematic and careful description of the task or workplace and has been planned to serve as a tool to help the analyst to form a foundation of the work situation [12]. To complete the tridimensional analysis, an observation tool was developed for the clients' dimension: questionnaire.

The ergonomic factors that allow operationally the ETdA model are the observable or measurable part of the model and are correctly identified in table 2.

Table 2. ETdA ergonomic factors

Environment ergonomic factors	Occupational ergonomic factors	Personals ergonomic factors
Noise	Professional training quality	Work postures
Illumination	Decision making	General physical activity
Thermal environment	Restrictiveness	Communication/inter-relation
Risk accident	Job content	Attentiveness
	Work space dimensions	

The authors consider that ergonomic factors can be intrinsically or whether be extrinsic linked to the professional, being respectively considered as environmental or occupational. If they are inserted in the organizational schemes of the social-technical systems it will be occupational, otherwise, environmental, when related to the involvement of the professional [11]. The properly study of the ergonomic factors in a

tridimensional perception will assess risk situations in commercial areas with free circulation of people and make easy the Ergonomic intervention.

2.1 Questionnaires

In ETdA model, questionnaire is the observation tool for client's dimension. It is a direct administration tool, who presents as a main advantage, the possibility of quantification of a variety of data and consequent establishment of multiple correlations. The questionnaire is divided into the following general sections: client's identification, client ergonomic analysis with an evaluation of environmental, personal and occupational parameters perceptible to the client, and open question [13]. The first part of the questionnaire asks about social-demographic characteristics like gender, age and professional activity. Clients' were also asked about the meaning of the ergonomic concerns and establishment preferences reasons. The questionnaire second part consists of client's perception of ergonomic issues, considering three major groups: environment, occupational factors and service quality. The ergonomic factors are identified in table 2. In the third part there is an open-ended question about the respondents' general satisfaction with the establishment. Respondents are asked to say in their own words what could be improved in the service provided and about the establishment general appearance. This issue can have high importance for management total quality, as it reveals client's perception and opinion of the establishment. The questionnaire was pre-tested to be used in the survey [14].

2.2 Evaluation forms

Professionals used the evaluation forms to do the assessment of the commercial area, adapted to this particularly economy sector [5]. The response form is a simple sheet on which professionals where asked to state about the different ergonomic factors (see table 2) related to all the CAFCP. The response set categories' vary in ascending order, according to the seriousness of the situation (Table 3).

Table 3. Ordering categories

Very bad	Bad	Good	Very good
--	-	+	++

2.3 Procedures

A total of 600 questionnaires were delivered (100 for each establishment) to bedistributed to clients and professionals. Each questionnaire was directly delivery in hands by the professionals and was completed in locus or at home. The professionals' assessment of the commercial area was done anonymously.

2.4 Results

The results shows that client's age ranged between 13 to 75 years old, with a mean age of 36 (Figure 2) and about 88% of the respondents were female. The existence of two clients aged 13 and 14 years, increase the liability of the establishment, reinforcing the importance of potential ergonomic risk characterization.

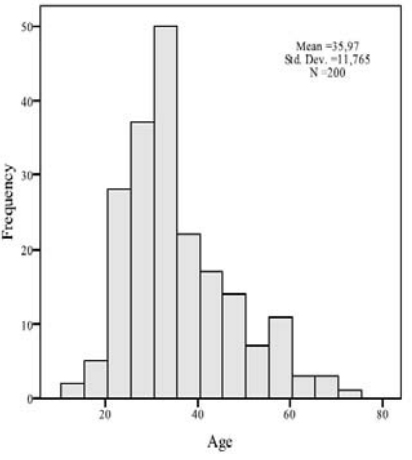


Fig. 2. Clients' age distribution

Clients and professionals results were submitted to statistical treatment and subject to reorganization for a combined analysis. The different answers categories' included in the questionnaire and evaluation forms were assigned according to three major groups: negative, satisfactory and positive categories. Table 4 presents the categories distributions.

Table 4. Answers' categories distribution

Categories		
Negative	Satisfactory	Positive
Always	rarely	Never
often	good	unlikely
very bad, bad	sometimes	very good
dissatisfied	reasonable	impossible
little		

In order to analyze with more detail the effect of clients' and professional evaluation in the ergonomic approach, a deeper analysis was done. Table 5 and 6 shows, respectively, a summary of the clients and professional result's.

Table 5. Clients' categories distribution (%)

	Categories		
	Negatives	*Satisfactory*	*Positive*
Noise	3.7		96.3
Quality illumination	0.5	9.6	89.9
Risk accident	5.23	82.72	12.04
Professional training quality	2,6	44.1	52.8
Anthropometric dimensions (trade desk)			76.8
Anthropometric limitations (shelf)	8.9	12.5	78.6
General physical activity	18.6	77	4.8
Restrictiveness	9.8		84.5

Table 6. Professionals' categories distribution (%)

	Categories		
	negatives	*Satisfactory*	*positive*
Noise	18.8	62.5	18.8
Quality illumination	6,3	62,5	31,3
Risk accident	13.4	73,3	13,3
Thermal environment	12,5	68,8	18,8
Work space dimensions(trade desk)	37.6	43.8	18.8
Work postures	18,8	62,5	18,8
Job content		71,4	28,6
Decision making	20	64,3	28,6
Lifting activities (shelf)		73,3	26,7
General physical activity		75	25
Restrictiveness	6,3	62.5	31.3
Attentiveness		56.3	46.8
Communication/inter-relation		40	60

The client's results show that the environment factors, noise and quality lightning, have a positive answer tendency, like anthropometric issues, quality service provide (professional training and restrictiveness). The risk accident perception and general physical activity have the largest percentage placed in satisfactory categories. The professionals' results follow satisfactory category tendency. The agreement in clients and professionals positive results has more impact in the proposals of changing's (ergonomic intervention) to be presented to the top management.

In the ETdA model, negative and satisfactory categories should not be underestimated. They can symbolize an indicator for a supplementary ergonomic study. Therefore, the existence of categories in the two dimensions, that show severity in every ergonomic factors analysed should be considered according to the relevance of each subsystems involved. To support this theory, the obtained results were also analyzed in order to understand the impact of negative and satisfactory categories. For example, in the client's evaluation the existence of negative categories in the professional training quality, sustains the idea that investment in professional training is beneficial

not only to the corporate profits but also to increase service quality and client's fidelity. Therefore, this particularly issue it is related with the subsystems: management strategies.

The analysis of the restrictiveness issue confirms the importance of the negative categories in clients' results in the ergonomic intervention goals. If the analysis were only performed by professionals, the 6.3% obtained could be devalued. Since this issue is related with the client's service quality, and this is a particularly top management principle, the 9.8% related to the client evaluation is an indicator for more and better investment in this area. Evidently, professionals will benefit of this procedure management. A similar analysis can be done in the other issues, showing that clients' involvement in the ergonomic analysis will benefit the ergonomic intervention and consequently the professional environment.

3 Final considerations

The development of new socio-technical systems in the market trade economy, leads to an ergonomic context where clients are participators, interacting with professionals in commercial common areas. A study case was done in the health sector to explore the possibility that using client's ergonomic evaluation will affect the ergonomic issues contributing for the changing proposals to be implemented. The findings revealed that if the analysis were only performed by professionals, some important issues could be devalued and consequently the ergonomic intervention would not be as effective as would be desired. The data analysis shows clients' contribute significantly to the final tridimensional analysis, since they highlight some risky situations that otherwise could be underestimated. It is easier to make organisational changes when the principal intervenient, client, as the same opinion then the analyst and/or the professional. This co-production of ergonomic list of priorities seams to be more effective, contributing to an efficient ergonomic intervention.

The proper study of the clients' involvement in ergonomic analysis is in line with the future challenges that propose a greater involvement of the public on ergonomic issues.

References

1. Querelle, L., Thibault, J.F.: The practice of the ergonomist consultant: a reflexive tools-based approach. @ctivités 4(1), 160–169 (2007), http://www.activites.org/v4n1/v4n1.pdf
2. Vink, P., Imada, A.S., Zink, K.J.: Defining stakeholder in involvement in participatory design processes. Applied Ergonomics 39, 519–526 (2008)
3. Virkkunen, J.: Le dévelopement collaboratif d'un nouveau concept pour une activité. @ctivités 4(2), 151–157 (2007), http://www.activites.org/v4n2/v4n2.pdf
4. Lindon, D., Lendrevie, J., Rodrigues, J., Dionisio, P.: Mercator XXI: Teoria e Prática do Marketing. Publicações D. Quixote, Lisboa (2000) (in Portuguese)
5. Taveira, A.D., James, C.A., Ben-Tzion, K., Sainfort, F.: Quality management and the work environment: an empirical investigation in a public sector organization. Applied Ergonomics 34, 281–291 (2003)

6. Carayon, P.: Human factors of complex sociotechnical systems. Applied Ergonomics 37, 525–553 (2006)
7. Loureiro, I., Leão, C.P., Arezes, P.M.: Tabela de ponderação: construção de uma metodologia para intervenção ergonómica. In: Arezes, P., Baptista, J.S., Barroso, M.P., Carneiro, P., Cordeiro, P., Costa, N., Melo, R., Miguel, A.S., Perestrelo, G.P. (eds.) Occupational Safety and Hygiene – SHO 2010, pp. 299–303 (2010) ISBN 978-972-99504-6-9
8. Loureiro, I.: Desenvolvimento de um Modelo de Avaliação Ergonómica em parafarmácias: Identificação e caracterização de pontos críticos e relacionamento com aspectos da população utilizadora. Thesis (MSc), Universidade do Minho (2008)
9. Loureiro, I., Leão, C.P., Arezes, P.: Modelo de Análise ErgonómicaTridimensional: impacto nas áreas comerciais com livre circulação de pessoas. In: Arezes, et al. (eds.) International Symposium on Occupational Safety and Hygiene - SHO 2009 Proceedings, pp. 273–277 (2009)
10. Ahonen, M., et al.: Ergonomic Workplace Analysis. Ergonomics Section Finnish Institute of Occupational Health, Finland (1989)
11. Loureiro, I., Leão, C.P., Arezes, P.: Desenvolvimento de um Modelo de Análise Ergonómica: impacto da população utilizadora na Análise. In: Arezes, et al. (eds.) International Symposium on Occu-pational Safety and Hygiene - SHO 2008 Proceedings, pp. 179–182 (2008)
12. Caple, D.: Emerging challenges to the ergonomics domain. Ergonomics 51(1), 49–54 (2008)
13. Hill, M.M., Hill, A.: Investigação por Questionário. Silabo (2000) (in Portuguese)
14. Khalid, H.M., Helander, M.G.: A framework for affective customer needs in product design. Theoretical Issues in Ergonomics Science 5(1), 27–42 (2004)

A Security Audit Framework to Manage Information System Security

Teresa Pereira[1] and Henrique Santos[2]

[1] Polytechnic Institute of Viana do Castelo
Superior School of Business Studies, Valença, Portugal
[2] University of Minho, School of Engineering
Information System Department, Guimares, Portugal
tpereira@esce.ipvc.pt, hsantos@dsi.uminho.pt
http://www.esce.ipvc.pt, http://www.dsi.uminho.pt

Abstract. The widespread adoption of information and communication technology have promoted an increase dependency of organizations in the performance of their Information Systems. As a result, adequate security procedures to properly manage information security must be established by the organizations, in order to protect their valued or critical resources from accidental or intentional attacks, and ensure their normal activity. A conceptual security framework to manage and audit Information System Security is proposed and discussed. The proposed framework intends to assist organizations firstly to understand what they precisely need to protect assets and what are their weaknesses (vulnerabilities), enabling to perform an adequate security management. Secondly, enabling a security audit framework to support the organization to assess the efficiency of the controls and policy adopted to prevent or mitigate attacks, threats and vulnerabilities, promoted by the advances of new technologies and new Internet-enabled services, that the organizations are subject of. The presented framework is based on a conceptual model approach, which contains the semantic description of the concepts defined in information security domain, based on the ISO/IEC_JCT1 standards.

Keywords: Information security management, information system security, audit information system, ontology and conceptual model.

1 Introduction

The rapid advances of the information and communication technologies, in particularly the Internet, and its increase use, have promoted the speed and accessibility of operations, resulting in significant changes in the way organizations conduct their activities. Consequently organizations become increasingly dependent on the availability, reliability and integrity of their information systems to be competitive and create new business opportunities [5]. However, the use of information technology brings significant risks to information systems and particularly to the critical resources, due to its own nature. An increased number of sophisticated attacks are expected to evolve as wireless and others technologies transcend. This fact enforces the

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 9–18, 2010.

need to ensure the security of the organizations information systems. In this context, it is crucial to perform a proper management of security, through a continuous identification of the main assets and their vulnerabilities, as well as the threats and attacks that they be subject of. One strategy to approach this goal is to perform regular information system security audits, to evaluate the performance of the security information management and analyze if the existing security practices needed to be reviewed. A security audit of an information system is conducted to assess the effectiveness of an organizations ability to protect its valued or critical assets [10]. This paper intends to present an investigated approach to improve security management through a conceptual framework developed to assist organizations to classify attacks, identify assets and mitigate their vulnerabilities and threats. The proposed framework is based on a conceptual model with capability to represent the semantic concepts and their relationships in the information security domain, defined accordingly to the established security standard ISO/IEC_JTC1[1] [8]. The paper is structured as follows: in the section 2 it will be presented an overview of security management concepts; in section 3 we briefly introduces the related work in information system security domain; section 4 presents the proposed conceptual model, which contains the semantic concepts specified in the information security domain, and their relationships, hierarchical structured in an ontology; section 5 presents the proposed framework to manage and audit information systems security, based on the ontology structure; conclusions and future work are presented in section 6.

2 Security Management

Managing information system security is increasingly concerning organizations, due to the continuous growing dependence of organizations on technology to conduct theirs businesses and to create a competitive advantage. Organizations rely significantly on technology, such as Internet, for businesses operations and secure business transactions [11]. Its recognized that since the last decade, the organizations are more dependent on the use of computer networks for their operations. Society as well as governments, depends on the use of computing services for their administration. Institutions depend on the effective use of computers and theirs communications for administrators to perform their daily operations, while the military requires secure communications to disseminate classified information. These fact turns computer networks a critical asset. However, beside computer network, organizations have others critical resources, accordingly to their structure, objectives and activities. This fact enforces the complexity of managing information system security for organization, due to the diversity of their assets and respective value, which requires to be properly protected [11]. Additionally, sophisticated attacks have been developed in order to exploit new vulnerabilities in the critical assets, when new technologies and new Internet-enabled services transcend. As a result, organizations need to evolve security management strategies in response to the evolving information security requirements. Nowadays, a properly security strategy demands for a rigorous process, similar to any

[1] International Organization for Standardization (ISO)/ International Electro technical Commission (IEC), Joint Technical Committee (JTC 1).

other business process, where every agent interacting with critical resources need to be aware and participate in security management, both adopting secure behaviors and continuous evaluating security control's performance [3]. The performance evaluation of regular information system security audit is one approach to evaluate the organizations information systems practices and operations. An auditing process will enable to obtain evidences whether the organizations information systems security policies, maintains the assets integrity, confidentiality and availability, and operating effectively and efficiently to achieve the organizations security objectives.

3 Related Work

Information system security auditing is a process that evolves with the security needs of the business activity of any organization [12]. Hence different approaches exist for conducting and managing security audits, to help the auditors to support a security auditing process. However most of the available models or frameworks to support the security audit are generally based on more or less liberal interpretations of the security fundamental concepts [10]. Lo and Marc-hand present a case study of security audit of a medium-size organization [9]. Their study focused exclusively on specific audit components, such as infrastructure, remote access and wireless LAN audits. Moreover Baharin et al., propose a third party security audit procedure that solely concentrates its study on a single data analyzer for security auditing, such as firewall log audit [1]. The ISO/IEC 15408 -Common Criteria (CC) introduce some models that relates threats and vulnerabilities, presenting security concepts and their relationships, in terms of organizations, safeguards, risks and assets [2]. However this model has some limitations, since it doesn't include the representation of the vulnerabilities in asset neither establish a relationship between the vulnerabilities to other security concepts that are essential in protecting assets, accordingly to the increasingly grow of attacks that exploit the assets vulnerabilities [12]. Notwithstanding its importance, the CCs model is more focused and useful in the evaluation of engineering products [12]. Farahmand et al. [4], propose a model to classify threats and evaluate the effectiveness of the associated controls, in order to identify possible outcomes of an attack. However the relationships to other security resources, such as vulnerabilities in assets are not fully covered [12]. Besides these contributions there are also available some guidelines for information security management systems auditing, such as the one released in 2007 by ISO/IEC[7]. The Information Security Audit and Control Association (ISACA[2]) also provides security guidelines for security audit processes, and SANS[3] (System Administration, Networking and Security) developed the ISO 17799 Checklist [6]. These standards precisely define the main procedures, but are limited concerning the strict relations or process flows necessary to undertake a security task, such as an audit. In this paper it is proposed a unified conceptual framework to support auditors to conduct a proper audit within and organization and hence to improve a better management of information systems security.

[2] http://www.isaca.org/
[3] http://www.sans.org

4 Proposed Conceptual Model Developed in the Context of Information System

The proposed framework is based on a conceptual model represented by an ontology. The adoption of an ontology structure was considered to be an appropriate strategy to organize and structure the fundamental terminology and concepts involved in the security information domain. The defined concepts are based on a wide recognized standard, produced by ISO/IEC_JCT1. The study of attacks, threats and the assets' vulnerabilities in an information system continues to grow because it is evolving and has significantly impacts on an organization. Managing those concepts requires both a detailed understanding of security concepts and their relationships. Such understanding can assist organizations in implementing the right combination of protection controls to mitigate security risks related with the assets' vulnerabilities. The implementation of a conceptual model, richly represent security concepts and their relationships in terms of threats, attacks, vulnerabilities, assets and countermeasures [12]. The advantages of this approach to organizations are that enables them to: (1) properly identify the valued or critical assets; (2) properly identify the vulnerabilities of assets; (3) identify and mitigate potential threats; (4) evaluate the risks; (5) evaluate the efficiency and effectiveness of the security policies and safeguards defined and therefore analyze and implement the necessary adjustments to security policy adopted. The proposed framework, based on conceptual ontology with capabilities to jointly model attacks, threats and vulnerabilities resources, and their relationships to other security concepts, stands an important advance in managing information systems security. The defined conceptual model comprises 8 concepts and 16 relationships, based on the security standards ISO/IEC_JCT1 and was represented on an ontology structure, as illustrated in the figure 1. These concepts are described as following:

Incident – A single or series of unwanted or unexpected events that might have significant probability to compromise the information system security.

(Security) Event – An identified occurrence of a particular set of circum stances that changed the status of the information system security.

Asset – Any resource that has value and importance to the owner of the organization, which includes information, programs, network and communications infrastructures, software, operating systems, data and people.

CIA – The information properties to be ensured, namely: confidentiality, integrity and availability; besides these main security properties, and depending on the context, other security properties may need to the addressed, such as: authenticity, accountability and reliability.

Threat – Represents the types of dangers against a given set of properties (security properties). The attributes defined in this concept follow the Pfleeger approach (Pfleeger 2007), which include an attacker actions or position to perform an interception, fabrication, modification and interruption, over a resource.

Attack – A sequence of actions executed by some agent (automatic or manual) that explore any vulnerability and produce one or more se curity events.

Control – A mechanisms used to detect an incident or an event, to protect an asset and their security properties, to reduce a threat and to detect or prevent the effects of an attack.

Vulnerability – Represents any weakness of the system.

In short, the rational behind the ontology is structured as following: an incident is made from – *madeFromEvent* – events; the occurrence of an event can leadtoa lostof – *lostOf* – a set of security properties (CIA); an asset has security properties – *hasSecurityProperties* – and each one can be *affected* by a threat; on the other hand, a threat can *affect* one or more security properties; and finally, an asset *has* vulnerabilities. A threat is *materialized* by an attack, while the attacks *exploit* one or more vulnerabilities; an attack is also triggered *toward* an asset. Further, the implementation of control

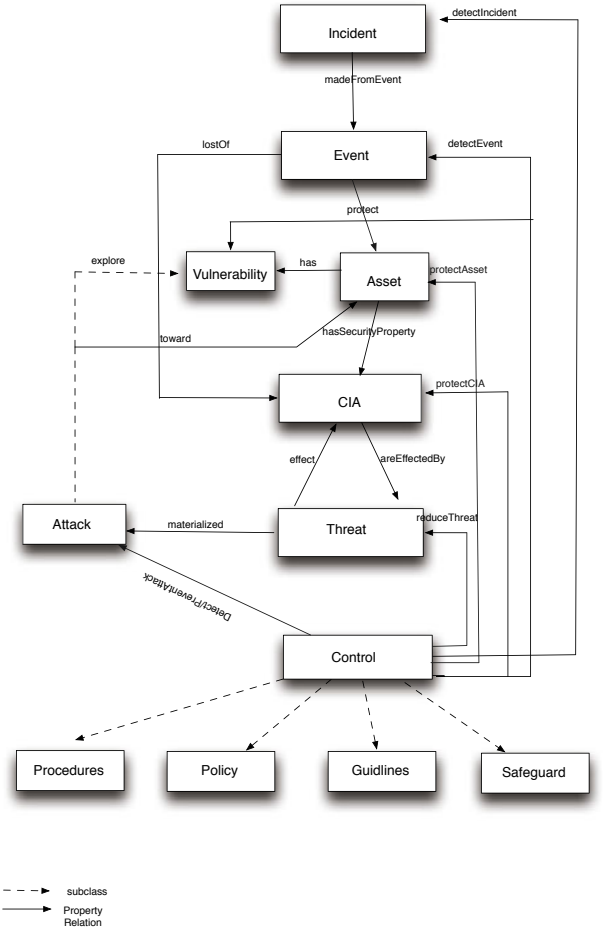


Fig. 1. Concepts and relationships defined in the conceptual framework

mechanisms, help to *reduc*e threats, to *detec*t and *preven*t an attack, to *protec*t security properties; to *protec*t assets and vulnerabilities, as well to *detec*t events, in order to *protec*t assets [13] . The description of those concepts and their relationships, presented in the ontology, was formalized through the use of the W3C standard language for modeling ontologies Web Ontology Language (OWL). This web language has been developed by the Web Ontology Working Group as a part of the W3C Semantic Web Activity [16]. In spite of OWL has not been designed to specifically express security issues, it was selected because it is a W3C recommendation since February of 2004 and due to its expressiveness with superior machine interpretability. The OWL is build upon Resource Description Framework (RDF) and Resource Description Framework Schema (RDFS). In fact the OWL vocabulary is an extension of RDF and uses RDF/XML syntax. The formalization of this ontology in OWL will be a step forward to promote its interoperability among different information security systems. In the next section, it will be presented the framework under proposal, which follows the hierarchical structure of the semantic concepts represented in the defined ontology, and try to provide an easy way to understand user interface so all users in an organization can participate in security auditing like tasks.

5 Proposed Framework to Manage and Audit Information System Security, Based on Ontology

The establishment of ISO/IEC_JTC1 standards promoted the standardization of the semantic concepts defined in the information security domain. The correct understanding and identification of those concepts are the primarily requirement to be considered in the performance of a proper examination of the information system security effectiveness, and further to identify and characterize an occurred security incident, as well as to estimate its impacts. The proposed conceptual framework intends to assists the organization, firstly to precisely determine what should be protected (the assets) and their weaknesses (vulnerabilities) involved in their daily activity. Secondly assess what vulnerabilities can be exploited by an attack, as well the threats that might be materialized in an attack. Finally, evaluate the efficiency and the effectiveness of the policy and controls implemented, in order to evaluate if they are being correctly implemented or if they need any adjustment [13]. Figure 2 illustrates the conceptual framework proposed, presenting these three nuclear concepts: attack, threat and assets. The auditor can select the concept from which he/she intends to start the auditing process, and proceed to the directed related concepts. Each concept contains a list of elements that are linked to the other concepts, conforming to the hierarchical structure of the semantic concepts, defined in the ontology. These three concepts were included in the front-end of the framework, rather the others, due to the nature of the audit operation, which the auditor intends to perform. Traditionally, a security audit is conducted once an incident has occurred (reactive followed by a corrective audit), that is when an asset has been compromised. In this case, an audit is requested in order to determine the source of the attack and how the incident happened, proceeding with the adequate corrective mechanisms. However a security audit is not only about investigating security break-ins, but rather to mitigate recognized threats, in order to ensure: (1) the security compliance; (2) the security of critical assets; (3) the right

controls are in the right place. In this last view a security audit is performed in the context of the security risk management process, and aims to produce or evaluate a security policy. Being conducted by the main concepts and their relationships defined by an ontology, the proposed framework intends to assist organizations to understand, prepare and perform security audits, by themselves. This framework does not focus exclusively on technical controls involved with information security, but enforces procedures and practices to assist organizations to maintain consistently high levels of useful and good quality information concerning their information security systems.

Within the ontology, each concept is mapped to real subjects. For example a malicious code attack, as illustrated in the figure 3, includes a brief description of its main features, followed by the available connection/link to the affected assets, the vulnerability it explores, and the security properties that have been compromised, as illustrated in the figure 4. Despite the large amount of information available to complete a basic ontology, we accept that each organization will develop its one view of security awareness. The framework is modular concerning this aspect, allowing evolving the ontology by adding the relevant subjects. This way, the auditor may proceed through the examination of the relevant vulnerabilities in the assets that can compromise the security of the information system, within the organization; or the auditor may go along with the analyses of new threats that might be materialized in an attack.

Additionally, the proposed framework includes the typical functions of similar tools, enabling a set of functionalities, like the possibility of the auditor to generate a report with all steps performed, as well as the registration date of the audit. According to the results of the auditor examinations, he can also schedule the next audit. Moreover, if the auditor during his examination detects a new incident, i.e. an attack that is

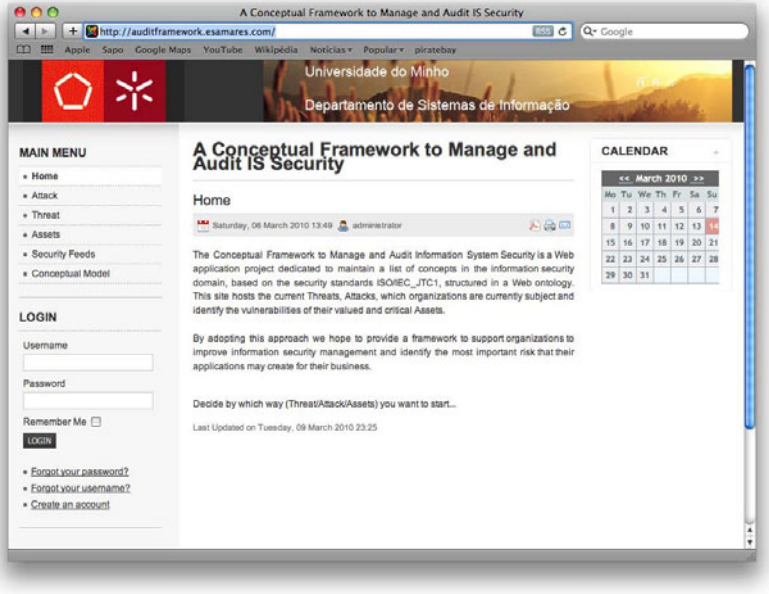


Fig. 2. Print screen of the developed framework

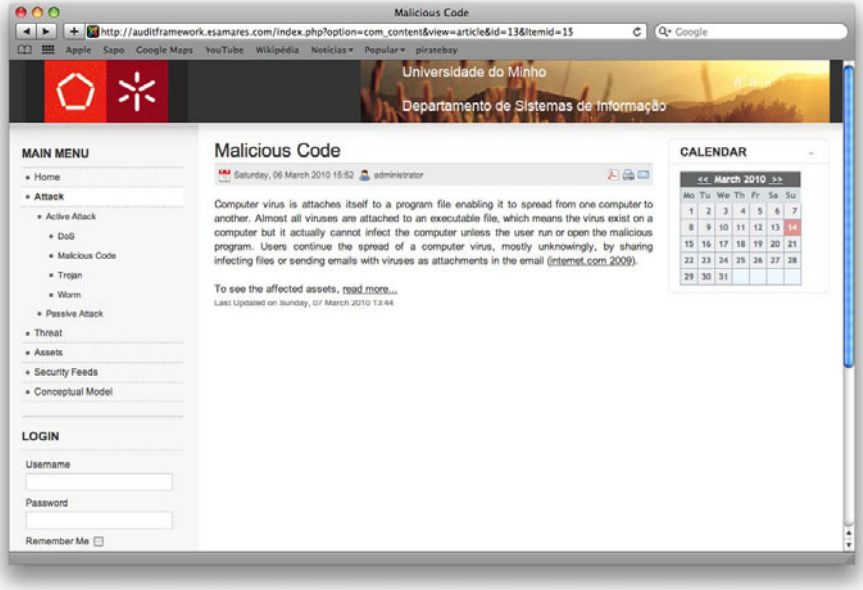


Fig. 3. Print screen of the developed framework

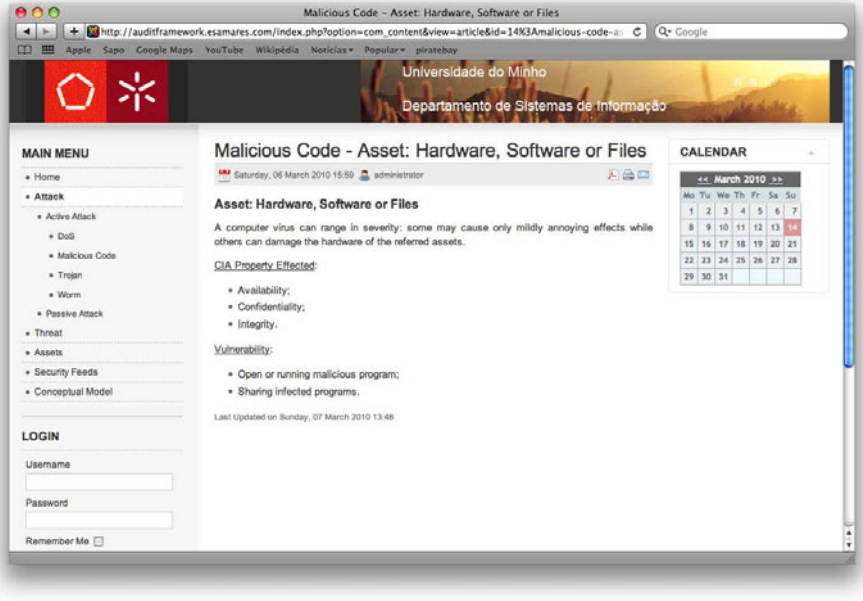


Fig. 4. Print screen of the assets affected by a Malicious Code attack with de security properties compromised and the vulnerabilities exploited by this attack

not presented on the list of attacks, the auditor should report this new attack with its features, which will be validated by the administrator of the framework and, after that, the administrator will index the attack to the list of attacks. This procedure is the same if the auditor decides to conduct the audit through the examination of the assets or threats and during the process identifies a new vulnerability in an asset or a new threat. The development of this framework is in a preliminary stage and it needs further improvements. However the concept of this framework is to assist the auditing process and promote improvements to the current methodologies available for information security management.

6 Conclusions and Future Work

Managing attacks and threats which the information system of the organizations are subject of, is increasingly difficult, due to the natural evolve of the attacks, threats and vulnerabilities, promoted by the advances of technology [5]. Auditing security information management is essential, and should not be performed only when an incident occurred, but also to assess if security controls and procedures are adequate and compliant with local or global regulations. The main contributions of this paper is a proposed framework based on a conceptual model approach, to support the auditor to primarily understand the business requirements in managing security of an organization, through the (1) exactly identification of the assets that needs to be protected; (2) identify and assess the vulnerabilities in the assets; (3) identify the potential threats that could be materialized in attacks; (4) evaluate the risks; (5) finally, assessment or reassessment of the policy and controls adopted. This solution introduces a new perspective to model information, in the security domain. Actually, a framework based on a conceptual model with capabilities to richly describe multiple security resources within an organization is an important advance, compared to the current models that typically address general purpose security issues. Besides the aforementioned advantage, its pertinent to highlight that it also promotes firming up and unifying the concepts and terminology defined in the scope of information security, based on the relevant ISO/IEC_JTC1 standards. Furthermore, it enables an organization evolving its own instantiation of the security ontology, obeying to standard concepts, but embedding its one view and assumed risk of exposition. As future work we intend to complete the implemented framework, introducing more elements to the concepts defined, as well as to implement the necessary adjustments to integrate further functionalities of the framework, e.g., direct link to attack and vulnerabilities description databases, alert mechanism for ontology outdate and continuous monitoring of security controls to promote early detection of security policies breaks.

References

1. Baharin, K.N., Md Din, N., Jamaludin, M., Md Tahir, N.: Third Party Security Audit Procedure for Network Environment. In: 4th National Conference on Telecommunication Technology, Shah Alam, Malaysia (2003)
2. Common Criteria for Information Technology Security Evaluation, Part I: Introduction and General Model, Version 3.1, Revision 1, CCMB-2006-09-001 (September 2006)

3. Da Veiga, A., Eloff, J.H.P.: An information security governance framework. Information Systems Management 24, 361–372 (2007)
4. Farahmand, F., Navathe, S.B., Sharp, G.P., Enslow, P.H.: Managing Vulnerabilities of Information System to Security Incidents. In: Proceedings of ICEC 2003, Pittsburg, PA. ACM, New York (2003) 1 58113-788-5/03/09.
5. Hayes, B.: Conducting a Security Audit: An Introductory Overview. Security Focus, http://www.securityfocus.com/infocus/1697 (accessed January 2010)
6. Information Security Management, Part2: Specification for Information security management systems AS/NZS 7799.2:2003 BS 7799.2:2002. (2003) SANS, http://www.sans.org/score/checklists/ISO_17799_checklist.pdf
7. ISO/IEC FDIS 27000 Information technology – Security techniques – Information security management systems Overview and vocabulary. ISO copyright office, Geneva, Switzerland (2009)
8. ISO/IEC FDIS 27001 Information technology – Security techniques – Information security management systems – Requirements. ISO copyright office, Geneva, Switzerland (2005)
9. Lo, E.C., Marchand, M.: Security Audit: A Case Study. In: Proceedings of the CCECE, Niagara Falls, 0-7803-8253-6/04. IEEE, Los Alamitos (May 2004)
10. Onwubiko, C.: A Security Audit Framework for Security Management in the Enterprise. In: Global Security, Safety, and Sustainability: 5th International Conference, ICGS3 2009, London, UK, September 1-2 (2009)
11. Onwubiko, C., Lenaghan, A.P.: Challenges and complexities of managing information security. Int. J. Electronic Security and Digital Forensic 2(3), 306–321 (2009)
12. Onwubiko, C., Lenaghan, A.P.: Managing Security Threats and Vulnerabilities for Small and Medium Enterprises. In: Proceeding of the 5th IEEE International Conference on Intelligence and Security Informatics, IEEE ISI 2007, New Brunswick, New Jersey, May 23-24 (2007)
13. Pereira, T., Santos, H.: An Ontology Based Approach To Information Security. In: Sartori, F., Sicilia, M.-A., Manouselis, N. (eds.) Communication in computer and Information Science, vol. XIII, 330 p. (2009) (Soft-cover); 3rd International Conference, Metadata and Semantics Research (MTSR 2009), Milan, Italy, September 30th -October, pp. 183–193. Springer, Heidelberg (2009) ISBN: 978-3642-04589-9
14. Pfleeger, C.P., Pfleeger, S.L.: Security in Computing, 4th edn. Prentice Hall PTR, Englewood Cliffs (2007)
15. Walker, D.M., Jones, R.L.: Management Planning Guide for Information Systems Security Auditing, special publication of the National State Auditors Association and the U.S. General Accounting Office, December 10 (2001), http://www.gao.gov/special.pubs/mgmtpln.pdf
16. Smith, M.K., Welty, C., McGuinness, D.L.: OWL Web Ontology Language Guide, W3C Recommendation. Technical report, W3C (February 10, 2004), http://www.w3.org/TR/owl-guide/

The Cloud's Core Virtual Infrastructure Security

Annette Tolnai and Sebastiaan von Solms

University of Johannesburg, Johannesburg, South Africa
atolnai@global.co.za, basievs@uj.ac.za

Abstract. Cloud service providers (CSPs) should institute the necessary security controls, including restricting physical and logical access to hypervisor and other forms of employed virtualization layers. To enact relevant security measures, the core elements communicating with the hypervisor need to be secured. A proposed security model will introduce some of the aspects that need to be secured in the virtual environment to ensure a secure and sound cloud computing environment. This paper will discuss the core aspects of the virtualized architecture explaining the security risks, including a discussion pertaining to the relevant security core concepts to mitigate the risks.

Keywords: Security model, cloud, virtual, security, virtualization security, resources, security countermeasures.

1 Introduction

Host security responsibilities in SaaS, PaaS and IaaS are the responsibility of the cloud service provider (CSP), who has to be concerned about protecting hosts from host-based security threats. Some virtualization security threats such as virtual machine escape, system configuration drift, insider threats and root kits by way of weak access control allow new threats to the hypervisor which carry into the public cloud computing environment.

"*The integrity and availability of the hypervisor are of utmost importance and are key elements to guarantee the integrity and availability of a public cloud built on a virtualized environment."*[1] A vulnerable hypervisor could expose all user domains to malicious insiders.

"*Since virtualization is very critical to the IaaS cloud architecture, any attack that could compromise the integrity of the compartments will be catastrophic to the entire customer base on that cloud."*[1] In a virtualization environment, hypervisors have more access, such as root level access in some systems, to hardware resources compared to that of typical applications. Attackers can gain access to the entire host computer where the base systems as well as multiple virtual machines reside. When control of the hypervisor is obtained, data as well as sensitive information can be accessed and redirected.

We start with a discussion of how virtualization fits together with the cloud to provide efficient services to end users. We will then move on to a discussion of virtualization security where the security model is introduced.

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 19–27, 2010.

2 Virtualization and the Cloud

In short, cloud computing is the ability to scale and virtualize resources over a network, usually the Internet, such that a service is provided to the end user. Almost all IT resources can be delivered as a cloud service: applications, compute power, storage capacity, networking, programming tools, even communications services and collaboration tools. [2] Cloud computing combines service orientation with distributed manageability and economies of scale from virtualization. Without virtualization, the cloud becomes very difficult to manage. "*Virtualization is so important for cloud computing because it is possible to simplify many aspects of computing.*" [3]

Virtualization decouples software from the hardware, making it important for the cloud. Virtualization can be applied very broadly to memory, networks, storage, hardware, operating systems and applications. Additionally, a cloud infrastructure is based on a service-oriented architecture (SOA).

2.1 Virtualization, SOA, and the Cloud

Fig. 1 depicts a simple overlay of virtualization and the cloud powering the SOA architecture. Fig. 1 represents the service-oriented architecture (SOA) (Fig. 1: A) supporting the virtualized desktop and application services (Fig. 1: B) that are created, run and managed (Fig. 1: C) on a virtualized infrastructure, platform and software service (Fig. 1: D). Here both the service consumer, i.e. the business application or the business orchestration script, and the service provider are candidates for hosting on virtual infrastructure.

The virtual machine can communicate directly with the virtual server software beneath it, exchanging information about resource allocation and preventing the need for a separate full operating system layer between the two.

We can think of cloud computing as being the next stage of development for virtualization. For the use of resources to be effective, full-service management platforms must be implemented so that resources are safe from all forms of risk. As in traditional systems, the virtualized computing environment must be protected: [3]

- The virtualized services offered must be secure
- The virtualized services must be backed up and recovered as though they're physical systems
- The resources need to have workload management, workflow, provisioning and load balancing at the foundation.

Without this level of oversight, virtualization won't deliver the cost savings that it promises. A brief introduction to virtualization and cloud computing had been presented. We will now have a look at what virtualization security risks can filter into the cloud.

3 Virtualization Security Risks in the Cloud

Some virtualization security risks such as virtual machine escape, system configuration drift, insider threats and root kits by way of weak access control allow new

threats to the hypervisor which carry into the public cloud computing environment. Risks such as virtual machine escape, system configuration drift, insider threats and root kits can potentially cause a security breach in the virtual environment. Some of the security countermeasures and recommendations are discussed in section 5. An overview of the virtual infrastructure resources is provided next.

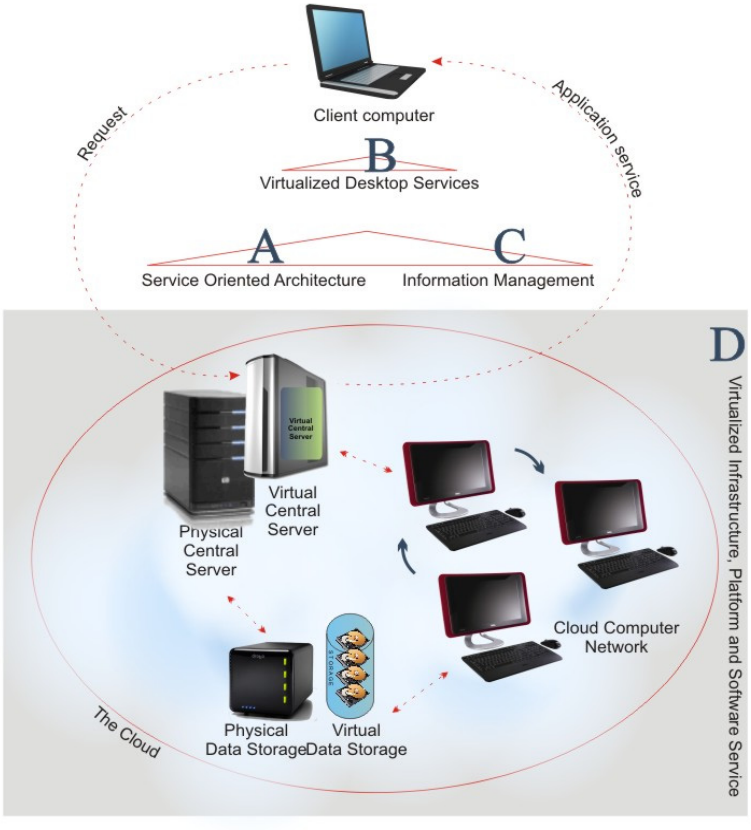


Fig. 1. A high-level depiction of virtualization, SOA and the cloud

4 Overview of the Virtual Infrastructure Resources

How the kernel governs access to external resources, and its security is important to understand. The virtual machine speaks to the virtual machine manager, which is a part of the virtual hardware layer, which speaks to the kernel on the virtual machine's behalf. The kernel regulates through the use of virtual switches the interaction between the guest, its virtual network interface card (NIC), and the outside world. The virtual switch is the connection point for kernel network devices, which connect to port groups on a virtual switch. The kernel must mitigate all indirect access to the hardware from each individual virtual machine through that virtual machine's virtual

machine manager layer. The virtual machine thinks it is talking directly to the hardware, but is really talking to the kernel, which talks to the hardware on its behalf.

The major resources are discussed below, which is accessed based on resource type by the virtual machine through the kernel. These resources create the greatest security risk [4] and contain the CPU, memory assignment, memory swapping, content-based page sharing (CBPS), access to network and access to disk, and APIs into the hypervisor. These elements will be discussed in the next section.

5 Security Countermeasures

The elements discussed will ensure the core of the virtual system is secure. The hypervisor sits between the hardware and virtual machines and several instances of may be in use within the virtual infrastructure. If the hypervisor is insecure, the rest of the computing environment is also insecure. If the hardware sitting below the hypervisor is insecure, the hypervisors and the virtual machines are also insecure. [4] "*Of course, the compromising of the hypervisor or the virtual middleware is one of the highest vulnerabilities and threats which are out there."*[5]

Although securing the virtual environment is not only about the technical aspect, other aspects such as management, audit, service level agreements (SLAs), trust and privacy amongst other non-technological aspects which contribute to security are beyond the scope of this paper.

The security risks mentioned in section 3 needs to be mitigated, amongst others. Although these security risks are not the only risks present in a virtual environment, there are relevant steps which can be taken to verify that security is well implemented against other security risks. These steps are as follows (see model illustrating the core virtualization infrastructure security elements in Fig. 2):

1 Secure the hardware
2 Secure the host operating system, if any
3 Secure the hypervisor
4 Secure the management interfaces
5 Secure the virtual machine

The model illustrating the core virtualization infrastructure security elements are a part of a bigger model, which also addresses above-mentioned aspects such as management. The purpose of this paper is to address the lower-level security elements. Each of the above steps will be explained further.

5.1 Secure the Hardware

A root kit can exist within the hardware by the act of patching firmware. The different types of known root kits have been discussed previously. The potential security solutions to detecting and patching the root kits are presented below: [4]

- Firmware root kits: In order to detect these types of root kits, the comparison of firmware checksums created using a cryptographically safe hash algorithm can be used by downloading the existing firmware, running a checksum and comparing it to a known good value. These types of root kits are difficult to find, and exist even after rebooting the system.

- Blue pill: There is a way that has been discovered to detect some of these root kits by looking at resource consumption of the translation look aside buffer (TLB). Concepts such as Guard Hype have been developed to aid in the prevention of these types of root kits, but there is no concrete product that is able to do this yet.
- Vitriol: This version sits within the hardware and is difficult to detect. The same detection mechanism as described for the blue pill root kit may be applied here.

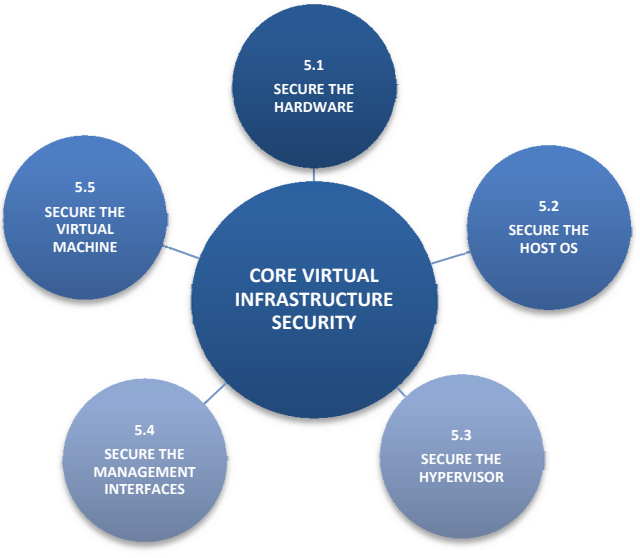


Fig. 2. Model illustrating core virtual infrastructure security

5.2 Secure the Host Operating System

A security assessment needs to be run to identify the security risks and to be able to fix the security gaps in order secure any management appliance. In order for the entire environment to be secure, the management appliance operating system must also be secure.

The hardening guidelines which are available assist in hardening the service do mainly cover the hardening of the service console or management appliance but other aspects of the virtual environment, such as virtual file system security and complete coverage of virtual network security are not covered.

5.3 Secure the Hypervisor

This step requires that all the levels of security within the hypervisor is understood. But because many hypervisors are proprietary, we cannot gain access in order to harden the security any further. However, the kernel is what interacts with the virtual machines and all the hardware. The virtual hardware with which the kernel interacts is the CPU, memory, network, disks and APIs.

5.3.1 Access to CPU

The central processing unit (CPU) is virtualized using virtual CPUs. More than one virtual CPU, which have a one-to-one mapping to physical or logical CPUs are shared by the scheduler. It is possible to assign a virtual machine to a number of logical CPUs based on the number of virtual CPUs within the virtual machine. The scheduler will use only the selected CPUs for running the virtual machine, which is done by setting CPU affinity. The security risk exists when utilizing CPU affinity as it is possible to produce a DoS style attack by assigning all virtual machines to the same CPU. This leaves most of the systems idle, but forces the scheduler to use only one CPU for all virtual machines. This would cause massive performance degradation where even the kernel would have issues, which could lead to data corruption and availability issues.

5.3.2 Memory Assignment

Whether a virtual machine is started up or memory is borrowed from another virtual machine, all memory is zeroed by the kernel when the balloon memory driver is employed. The balloon driver controls whether a virtual machine's memory can temporarily be borrowed from another virtual machine that is not currently using it, on an as-need basis. Unless there is memory constraints imposed on the memory allocation of the virtual machines, the balloon driver is not in use. The memory taken from one virtual machine is zeroed out before handing it over to another virtual machine, meaning that there is no security risk. Through the use of the kernel's memory management, it is not possible for one virtual machine to see another virtual machine's memory.

5.3.3 Memory Swapping

A security risks exists where the virtualization server administrator can access the virtual machine's virtual memory swap file and access its data, a snapshot can make a copy of the virtual machine's memory for later restoration, and any sleep mode used for the virtual machine. The copies of memory can be accessed by the super user of the virtualization host. To mitigate the security risk, it is important not to allow anyone to log in directly as super user. The built-in auditing functionality can be used to ensure files are not directly accessed or moved from the specific system.

5.3.4 Content-Based Page Sharing (CBPS)

CBPS enables virtual machines to share the same memory pages between themselves. To determine whether a page already exists within the CBPS memory area, all allocated memory pages are scanned and a hash is created of the page. If any of the hashes match, the CBPS implementation does a bit-by-bit comparison to determine if the page of memory is identical. Hash algorithms are no longer cryptographically safe, but the CBPS is safe, due to its bit-by-bit comparison. The kernel uses copy on write (COW) to detect changes to CBPS pages of memory, where a newly allocated page is created, with the existing page copied to the new page. The copied page is then written to while this prevents the CBPS pages from being modified. CBPS can also be disabled, but there is no security reason to do so. CBPS is enabled by default; page sharing can reduce the overall memory consumption of like virtual machines. For example, this can produce good results on an ESX Server, and is a current best practice.

5.3.5 Access to Network

The virtual switch increases functionality and some security, because it acts a simple software device that represents the Layer 2 network switch. When connecting virtual machines through virtual NICs, the virtual switch is the connection point for kernel network devices, which connect to port groups on a virtual switch. Network access is given to other kernel elements that run specific services. These devices don't have any built-in firewalls, but the protection is available through the virtual switch. It is possible to add firewalls into virtual switches, but currently a virtual switch will provide protection from the following types of attacks: [4] MAC flooding, double encapsulation attacks, multicast brute force attacks, spanning tree attacks, random frame attacks, ARP cache poisoning. When increasing protection on a virtual switch, three other settings can be set:

1 Prevent a virtual machine from spoofing a MAC address.
2 Drop outbound packets where the source MAC does not match the one currently set on the virtual NIC.
3 Deny the capability of a virtual switch to enter promiscuous mode for a virtual NIC attached to a virtual switch or port group.

5.3.6 Access to Disk

There are several virtual disk concepts which contain the contents of the guest operating system**.** Each file has security concerns as the files are created by the guest operating system, as well as the style of the disk may too include several concerns. Complete data access is permitted to the virtual disk files when granted super user access, while regular users of the system do not have this access. A major security issue is to ensure that backup tools create files with the proper permissions. World-readable files are created by various backup tools, and full disk backups are produced. Anyone who has access to the storage device can access this information.

Forensic scientists and even hackers make use of slack or unallocated space represent artifacts of previously written data, such as credit card details, evidence leading to conviction and business crucial information. If that section of the data store was previously used, it is possible that the virtual machine disk could contain old data. Based on allocation methods and whether the LUN is first zeroed, will determine the security risks regarding the various disk types. The zeroed thick disk and the eager zeroed thick disk options are the ones which should be the primarily used disk formats, as they are the most secure.

5.3.7 Application Programming Interfaces (API) into the Hypervisor

Research and development processes may require programmers to utilize APIs into the kernel. The concept of virtual appliances and APIs create more attack points into the kernel than previously. The virtual appliances can also be attacked if they are network aware. By utilizing digitally signed virtual appliances by the appropriate certificate authority (CA), these attack points can be controlled. Also special networking configurations need to take place to further protect the kernel during runtime.

5.4 Secure the Management Interfaces

The management interface consists of the management of the entire virtual infrastructure, a specific host, or the virtual machines. The entire virtual infrastructure and a

single host fall under virtual infrastructure management. Virtual machines fall under virtual machine management. [4]

5.4.1 Virtual infrastructure Management

The multifarious effect means that several different passwords are required for the management appliance and each of these servers. Instead of creating user roles and permissions, groups should be created. This entails that the management of the varied management appliances are simplified to the directory service or group in use. Additionally, a single administrator group should be created that is allowed direct access to the management appliance for debugging and determining problems. Direct access to the management appliance should be denied where all other groups are concerned.

It is important not to change the default protections for the super user account that deny direct access through SSH using that account. This is important in order to guarantee some form of auditing capability.

It is important to understand that intermediary programs will perform access to the kernel using a private API. Direct access to any management appliance exposes a huge amount of data that can lead to future attacks. The kernel leaves quite a few artifacts that management appliance can access, even though it is protected from abuse. These artifacts, in most cases, are disallowed by security policies used in the most security conscious settings.

5.4.2 Virtual Machine Management

Virtual infrastructure management tools can directly affect a virtual machine, for example, the virtual machine is unaware of what is happening if the administrator chooses to detach or attach a network to a virtual machine. A set of tools and drivers can be installed so that the virtual machine knows a little more about its environment.

The tools and drivers in use can affect the security of the virtual machine, for instance the various isolation settings that will improve the security of the virtual machine through external management appliances. The information seen by the management appliances or the functionality of some of the remote console tools can be decreased by some of these isolation settings. The console of a virtual machine can be accessed remotely through the use of graphical management interface, exposing some functionality that could be considered a security risk. The use of isolation tools is necessary to limit information leakage, to protect the innocent and to be compliant with standards and guidelines.

It is not impossible for a virtual machine to see a disk of another virtual machine as this all depends on the file sharing setup within the virtual machine, guest operating system clustering technology, or the way in which the virtual machine was originally created. These concepts would apply even when using physical machines.

5.5 Secure the Virtual Machine

Hardening and security guidelines expressed by each guest operating system should be applied independently of the virtual infrastructure. Currently it is impossible to deny that a virtual machine is in fact a virtual machine, and a virtual machine can be detected by looking at the MAC address, or the hardware in use. The virtual machine hardening should include steps so that the remote console has limited exposure and is

only used on a need-to basis. Because access to the console can grant access to other mechanisms to breach a system, no one should need to access the virtual machine console. Additionally, the hardening guidelines for guest operating systems should include steps to protect the remote console as well as any additional files added to the system.

6 Conclusion

An overall explanation of the kernel and the various security implications of it use have been presented. The various security aspects of the virtual environment, including the virtual machines, management interfaces, hypervisors, operating systems and hardware have been discussed representative of the model. The groundwork leading to a secure virtual environment has been placed.

Aspects such as SOA, virtualization and cloud computing have been discussed, illustrating the importance virtualization plays in an efficient delivery of services.

References

1. Mather, T., Kumaraswamy, S., Latif, S.: Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance. O'Reilly Media, Inc., Sebastopol (2009)
2. Sun Microsystems: Take your business to a higher level. Sun Microsystems Inc. (2009)
3. Hurwitz, J., Bloor, R., Kaufman, M., Halper, F.: Cloud Computing: For Dummies. Wiley Publishing, Inc., Indianapolis (2010)
4. Haletky, E.L.: VMware vSphere and Virtual Infrastructure Security: Securing the Virtual Environment. Rights and Contracts Department. Pearson Education, Inc., Boston (2009)
5. Kissman, U.: IT Virtualization - New challanges for IT Security Management: IBM (2009), http://tv.computerworld.hu/video/it-virtualization-new-challanges-for-itsecurity-management (accessed: March 10, 2010)
6. BEA and VMware: SOA and Virtualization: How Do They Fit Together? Whitepaper (2007)

Collaboration and Command Tools for Crises Management

Tapio Saarelainen and Jorma Jormakka

National Defence College, P.O. Box 07 00861 Helsinki, Finland
tapio.saarelainen@mil.fi, jorma.jormakka@vtt.fi

Abstract. Present collaboration tools are not suitable for the command process. The Collaboration tools, in the use of military entities, lack the elements for command and control and scheduling of resources. Available collaboration tools concentrate on uncontrolled distribution of information. They represent Situational Awareness (SA) tools for the cooperating entities, not the required solution for traceable and sophisticated Command and Control tools (C2-tools) applicable for the Crises Management Environment (CME) and Military Environment (ME). This paper presents tools for crises management, which enables the use of effective C2-tools, functioning along with the Resource Manager (RM) and scheduler. Given tasks need to be traceable afterwards for various purposes. On the base of collected data from the events, actions and reliability of different collaborating entities, a trustworth database of the characteristics of each entity can be formulated and utilized afterwards as a base knowledge on the collaborating entity. Collected data remains in the information repository and the collected data is used for identification purposes of an entity. C2-tools in ME and CME are systems of systems based on trusted entities, which will execute the commanded tasks reliably and in a given time. Reporting tools are out of the scope of this paper.

Keywords: Collaboration, Command and Control tools (C2-tools), Crises Management (CM), Resource Manager (RM).

1 Introduction

Paper is introducing a suggestion to solve the problem how to boost and enhance the collaboration tools tailored for crises management. This is done by presenting answers for the recognized dilemmas. First key question of this paper is: Where is the division through organization boundaries and how to merge the organizations? Second one is how to connect to the existing network system and utilize the pre-defined collaboration tools and how to ensure the controlled process of command and control. The paper presents a solution for the questions by introducing the RM, Business Processes (BP) and the scheduler and by identifying the division between the collaborating organizations.

The overall increasing interest in collaboration tools in the Military Environment (ME) has lead to a considerable growth of various types of applications and programs. Examples of these are Multi-National Experiment 5 (MNE 5) and various Maritime

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 28–38, 2010.

Situational Awareness (MSA) solutions, which are in an experimental stage. These tools can be recognized as SA tools between civilian authorities and military entities, not tools for command and control, which benefit from the Business Process. Present tools lack the Resource Manager, scheduler and information repository, which can be seen the potential key elements for the Military SOA (MSOA). To merge collaboration tools developed into different infrastructures and network environments including self-organizing virtual network links, appropriate collaborating tools, programs and available network systems. If collaboration does not exist between various entities, there will be a chaos and several needless accidents that can be prevented by the means of effective, continuous and trustworthy co-operation.

It is also a matter of a trust, or lack of it. The ME is similar to CME. Presented C2-tools are combinations of different communication solutions, which vary along with the technical evolution, therefore they remain unidentified. However, that is not the key point, whereas the concepts of trust, authentication, authority and continuous process of data collecting for each entity are the essential functions. Without commitment into applied rules and policies, the cooperation with an entity is impossible. All of the described processes are continuously logged can be traced afterwards. Each entity is responsible to execute the given tasks in a given time and at the same forward critical SA data to the system to ensure the efficiency and avoid fratricide.

2 Collaboration Tools

All entitities need collaboration for their mission success and survivability in all military and crises management operations. If an entity is not collaborating, it is taking a calculated risk to fail. Collaboration needs suitable tools and reliable and ubiquitous network system. If there is no continuous and reliable collaboration, there will be chaos, accidents and messy situations and unnecessary paralyzing.

Since the pace of war is increasing with Net-Centric Warfare (NCW), so are risks. If external entity appears in the theatre and doesnt collaborate with the military entity, there will be a risk to get under friendly fire and on the front of maneuvres, risking their existence. There is a need to automate command and control tools utilized in military and crises management dure to the increased tempo of operations. Computers equipped with the solutions presented in this paper are able to keep up with the tempo of events. Processes including metadata need to processed, analyzed, verified and finally stored for future purposes. Metadata is also needed to link the processes into correct events and control and command chains utilized by the RM and the scheduler.

Collaboration is the key issue in the battlefield among the allies, i.e. the Blue Force. Nowadays the White Force, i.e., Non-Governmental Organizations (NGOs), is an entity co-operating with the military. The NGOs are a security risk for themselves without collaboration with the military. Their existence is depending on reliable SA. The NGOs have to choose their co-operation entity (Red Force or Blue Force) to gain information from the operational area and to survive. Once the side has been chosen, the entity can evaluate risks and benefits if it connects the NGO into its network [2]. It is a question of bargaining and benefitting. Moreover, it is a matter of trust and operational security in the theatre.

Collaboration tools can be seen as computerized location programs, mapping tools, Voice over Internet Protocol (VoIP) tools, security data of various type, tools of social media, raw and unanalyzed critical data in data warehouses, sensitive data gathered and analyzed data in a data warehouses, raw video stream or a still frame, even a text message can be used as a collaboration tool, to name a few [10] [13] [14]. Besides the mentioned ones, modern technology enables the creation and evolving of new types of tools, which vary due to the requirements. It is impossible to predict the emergency rate and type of new collaboration tools. The nature and function of collaboration tools evolve along the technology and social behaviour of humans at an unpredicted pace and rhythm.

3 SOA Based System and the Resource Manager (RM)

The civilian entity's challenge in a task-centric environment is similiar to that of a Battlefield Commander, who depends on C4I2SR tools to perform rapidly in given tasks. To merge military and civilian entities (civilian authority or a member of a humanitarian entity) common tools, databases and access methods to the data have to be used in a commonly accepted and trustworthy manner. Since operational C4I2RS are today often based on Service-Oriented Architecture (SOA), SOA must be addressed with the issues of information sharing regardless of where the precisely needed data are stored [29]. At the moment, SOA lacks the introduced element; the Resource Manager (RM). When the RM is properly adopted and utilized, it will improve the efficiency of the SOA and thus ensure the best outcome of the shared resources used. In the presented solution the RM is not an encapsulated system as part of each resource as presented in [31], since the RM is a functional module to control and distribute the allocated resources, differing from what is described in [32].

Before the SOA or the RM can be utilized in the tasking processes, the tasking process has to be defined, understood, configured and the entities have to commit themselves into the processes and protocols of the tasking and SA [6]. The process to send a task is basically as follows: Task/request of information is defined (the authority to command/accept the given tasks, is the commander of the Area of Operation, AOR), task is given (in pre-defined format, in pre-defined format), and task is sent via available network system in the AOR (Internet, WPSN, Hubs, radios, and computers) [10], not to forget a self-organizing virtual network links and their components [17]. In the other end, where the task is received, the process is as follows: Incoming tasks are filtered via various subsystems, which are presented in Figure 1. Task is passed via authentication, via security policy, via acceptance policy, via willingness and understanding policies and finally via execution policy. After acceptance policy (task is accepted into the system) task is forwarded to RM, which verifies the task and commands it to the proper and available entity which is responsible for its execution. Entity becomes a task performer, responsible for task execution.

Benefits to achieve reliable and constant collaboration require trust between entities, coordination on security issues, authority and authentication. Technical improvements are in trust (competence, common data services), when commanded entity is willing and capable of doing the commanded service. In security issues tagging information and protection can be recognized as company policies. Once the task is understood (language, culture), the entity is capable of executing the command. Since authority is recognized, the entity is allowed to perform the given task.

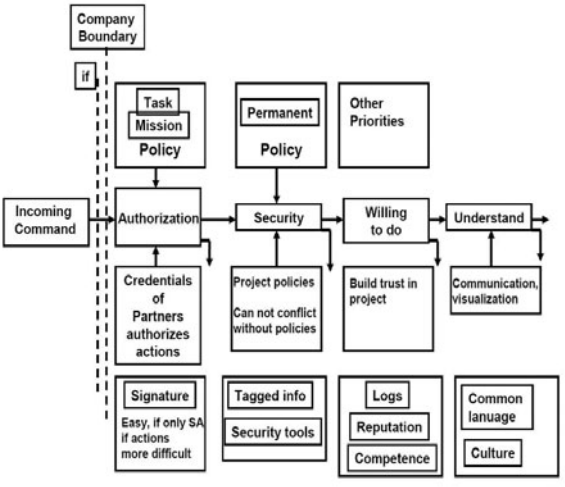


Fig. 1. A simplified sequence chart of command and task processing

Each task has a time-stamp and own identification id, it also contains route data, so it is to be tracked whenever tracking data is required. Each task will be categorized into urgency class and its execution process is been monitored and evaluated constantly. If the entity is incapable to perform the task, it can be retrieved or forwarded to new entity, responsible for its execution. The latter process can be executed only with the acceptance of the commander of the AOR. Once the task has been executed, it will be filed as a completed task into the common database. Tracking data of the completed task can be retrieved for analysing purposes at any time by the system operator.

Security, authentication and agreement tool is implemented in RM. Before any tasks are given to be executed or resources are given for use, the task or resource request goes via the described system, which is presented below in Figure 2. Incoming task passes through a preliminary phase, where it is checked and identified. Once task has been verified and approved and sent from a trusted and secure co-operation entity, it will flow via series of approval and authorization policies. The process ends with a phase where common language and tools are selected and then it moves forward inside the RM.

The overall description of the whole concept consists of three major parts and functions: 1) SA comprehending the existing solutions and tools, 2) C2tools and 3) Information repository. These three together enable the C2 process and saving of log-data for further analyses. These functions need the RM along with the scheduler to share and distribute the tasks and resources.

The RM sorts out and lines up the surfacing simultaneous requests concerning the demanded service or data. The RM shall be well protected against enemy actions. The RM serves as the element providing the needed services for User Groups (UGs) and can be either pre-programmed on demand or request basis [9]. User Groups (UGs) send a request for the demanded service, UGs are authenticated, UGs privileges are verified, and after these processes the request is transmitted to the RM. The defined actions of the RM are described in [33]. The RM is described in Figure 3.

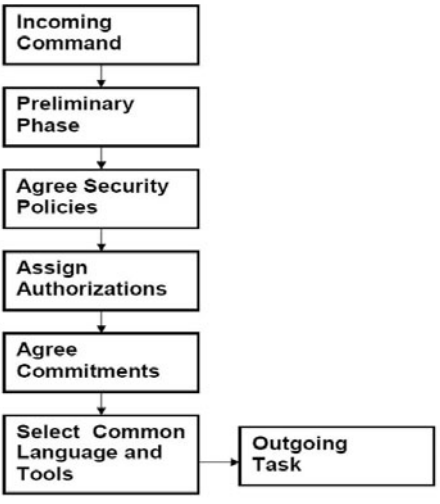


Fig. 2. Security, authentication and agreement system

Once the operation at AOR has lasted for long enough, substantial amount of data has been collected into the common database. Data comprises the efficiency of the each entity (time-stamps), amount of the tasks they have performed and the overall capability, efficiency and trust-value can be given to each entity. As co-operation goes on, more data is collected from each entity. Entities can be compared and verified on the base of given missions/tasks. Entities can be categorized into the different classifications based on their success in mission execution/task completion. The entities get credentials and increase their trustworthy by their reputation in mission success. On the other hand, organizations/entities collect credentials along the amount of essential data they have inputted into the system. Also the matter of co-operation/assistance in task execution and missions shall not be neglected while categorizing entities based on their earlier reputation.

The entities which work in the AOR, produce additional extended value into the system, utilize the collected data and enrich the common database by inputs of gathered information. This type of action is also appreciated whilst data from entity and its credentials are collected.

The question to be answered in this paper: Where is the division through organization boundaries and how to merge the organizations? Interfaces can be recognized as described in the Figure 4. The picture describes the division boundaries of the processes. Organizational boundaries can be merged via connecting them into same processes by forcing them to use the interfaces described below. Naturally, various entities are to be combined by offering them common collaboration tools (command and control tools) and access to the common databases. This requires commitment to obey common rules and policies consisting overall working and security principles.

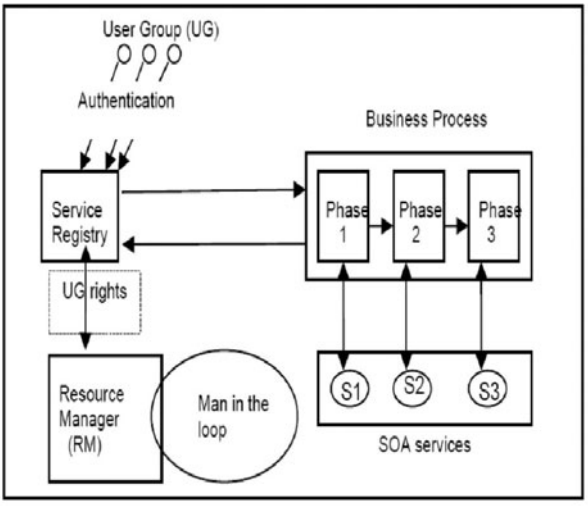


Fig. 3. The Resource Manager (RM)

4 SOA and RM

An SOA approach requires that business processes are represented as services. This creates a demand for a Business Process (BP) to be precisely defined before an assessment is made as to whether and how these can be implemented as services within a SOA. Service modeling of the business process is a critical step for creating meaningful services. If and when SOA is implemented as an enabler of BPs available for SOA, rewriting all existing capability is unnecessary. Techniques such as Service-orientation Migration and Reuse Technique (SMART) can be utilized to assist this process [24].

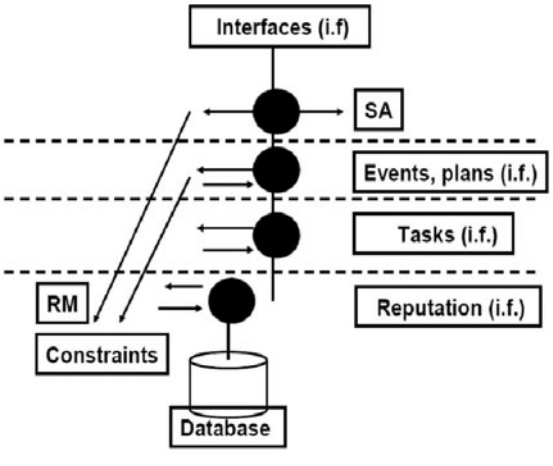


Fig. 4. Organizational boundaries described in command and control process

The Resource Manager (RM) communicates with four inter modules. The RM GUI provides the core interface between all the presented modules and the local area network, shown in the Figure 5. Local Area Network (LAN) is used as a battlefield network or a community network as it can be used on Wide Area Networks (WANs). The problematic issue remains, sharing of networking environment and its resources. Searching information and asking for resources is challenging without proper searching mechanisms. Each module has pre-defined and precise functions. First, the File and Resources Sharing module communicates with the RM GUI in conjunction with the sharing and the download module. The File and Resource Transfer and Download module supports and enables the transfer or download of the searched file or resource from the other node connected to the network. The shared files and resources are listed into the RM GUI, where the listed and downloaded files can be examined.

It is obvious that same services are requested simultaneously. Therefore the composition of the RM needs to be stable and reliable. Since the connectivity is a critical issue, it is essential to maintain the connection alive between the entities by securing a network connection with alternative routing. This problem is beyond of the function of the RM.

A critical component for MSOA is a scheduler. It schedules processes to avoid fratricide and collateral damage. The idea of the scheduler is to enable collaborating entities to carry out various operations simultaneously but under strict command and control. The issue of simultaneously operations is solved by the element named Battlefield Secure Scheduler (BSS). This component uses two different methods of sharing calendar Pre-Shared Schedule (PSS) and Dynamic Schedule Update (DSU). The scheduler functions along with the RM and utilizing MSOA as a process. These elements can be recognized in Figure 6, which introduces the process from incoming command/task to an outgoing command/task.

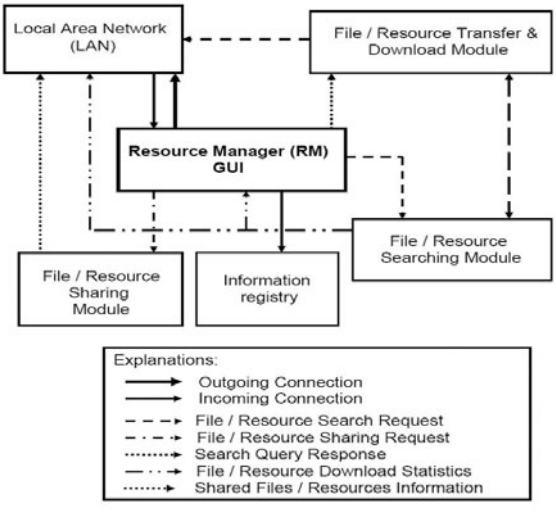


Fig. 5. The composition and function of the RM

5 Challenges with the Current Collaboration Tools

The usage and utilization of collaboration tools is based on self-interest to obtain the goals of the interest of entity‘s own organization and the entity itself. Therefore all co-operating entities have to benefit from each other in a convenient manner. Without the latter, co-operation with available collaboration tools does not happen.

Since the security issues are in the centre of the gravity, the risks have to be solved. This requires the commitment from the each entity to obey the rules agreed upon. An entity not committing into the processes and rules shall be eliminated from the co-operation. Abandoned entity is out of security, out of the help of collaborating entities, relying on its own information.

Utilization of collaboration tools is relying on stable and available network systems. Once there is no network connection, the collaboration tools are out of reach. The limited bandwidth of prevailing network in remote areas remains also a challenge [11]. The live video stream cannot be transmitted or received for the unstable connection and varying bandwidth [13] [14]. Also, there is a considerable challenge for the end-user on-The-Move (OTM) compared to the counterpart of At-the-Halt (ATH). To guarantee the required services for the mobile and stationary end-user, UVs (Unmanned Vehicles) can be seen as possible solutions if they are equipped with adequate power sources and directional antennas. Satellite communications can be also utilized when the equipment supports this possibility.

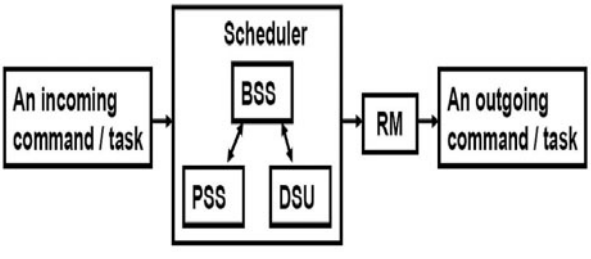


Fig. 6. The elements inside the scheduler and the permeable C2-process

6 Related Work

Multinational Experiment 6 (MNE 6), Multi-national Inter-agency Situational Awareness of the Extended Maritime environment (MISA-EM) and Situational Awareness (MSA) programs represent related work in collaboration tools in the case where military and civilian entities are co-operating. In various militaries, ongoing projects focusing on varied Net-Centric Solutions (NCS) to improve the overall capability and efficiency of national Future Armed Forces [25]. The following discusses two related projects, one from Australia and the other from the United Kingdom. The Australian Defence Force (ADF) is committed to transitioning, over time, to Network Centric Warfare (NCW), where NCW sees the elements of the ADF as nodes in a network [8] [26]. The UK Ministry of Defence (MoD) aims to significantly enhance military efficiency via the Network Enabled Capability (NEC) [27][28]. The NEC as a system

requires system integration of independent components, systems, and networks that can evolve, operate in a collaborative and dependable manner and manage system and component changes [8] [27][28]. In the case of White Force Tracking [3] the main goal is to avoid fratricide. In [30] new network system, Wireless Polling Sensor Network (WPSN), has been examined for the Dismounted Future Warrior (DFW) to improve their overall efficiency and performance in altering conditions and missions [4]. In [2] the possibilities of the DFW in system optimization have been evaluated.

Common object for the projects is the challenge of increasing the SA and information sharing and SA utilization in the training processes [6] [15]. Bearing the future battlefield in mind, to share the common tasks with an effective manner whilst the resources are diminishing, is the key into the success. The contribution of this paper is to present a solution how to combine the resources and how to share the data and given tasks in a common database with the selected and approved entities representing civilian and military actors. Certification and keying policies are essential along with the ad-hoc networks [5] in the utilization of collaboration tools [12].

7 Conclusions

This paper offers three results as a contribution for further development of C2tools. Results can be identified as 1) Command and Control tool (C2-tool), which enables the BP in the command and control process; 2) The RM, which is a central element of the MSOA in distributing of limited resources; lastly 3) The BP in the ME along with the MSOA. These results offer the missing attributes for the C2-tools for ME and CME. Combining these elements enable successful control for the BP in the ME and CME. Paper introduces the composition of the RM and the role of a scheduler, the function of the BP, and highlights the significance of trust and commitment in CME. If there is no trust, C2-tools are useless no matter what accessories and computer assisted tools are utilized. Trust is also needed to gather information of entities and to ensure tasks will be completed in the given time and in the given manner.

The question of benefitting from C2-tools is not only a matter of complicated command and control system, it is a matter of trust into the entity. Each entity embedded into the C2-tool environment can add increased value into the SA and intensify the outcome by committing themselves into rules and policies and obeying them. This way C2-tools can be benefitted from, otherwise they remain useless, increasing the amount of promising instruments for ME and CME.

References

1. Takahashi, Y., Sugiyama, K., Ohsaki, et al.: IEEE Group-Oriented Communication (2007)
2. Saarelainen, T., Jormakka, J.: Computer-aided Warriors for Future Battlefields. ECIW 2009 Press, Lisbon, Portugal, pp. 224–233 (2009)
3. Saarelainen, T.: White Force Tracking. ECIW 2009 Press, Lisbon, Portugal, pp. 216–223 (2009)
4. Feng, L., Shuguang, L., Lei, Z., Lou, L., Cheng, X.: Dynamic Virtual Prototype of Soldier System, Control Conference, IEEE (2008)

5. Jormakka, J., Jormakka, H.: A Certificate Revocation Scheme of a Mobile Ad Hoc Network. In: 8th Intl. Symposium on System and Information Security, SSI 2006 (2006)
6. Lampton, R., Cohn, J., Endsley, M., Freeman, J., Gately, M., Martin, G.: Measuring Situation Awareness for Dismounted Infantry Squads. In: Interservice/Industry Training, Simulation, and Education Conference, I/ITSEC (2005)
7. Boldstadt, C., Endsley, M.: Tools for Supporting Team SA and Collaboration in Army Operations, Collaborative Technology Alliances Conference (2003)
8. Khan, S., Tokarchuk, L.N.: Interest-Based Self Organization in Group-Structured P2P Networks, Consumer Communications and Networking Conference, IEEE (2009)
9. Takahashi, Y., Sugiyama, K., Ohsaki, H., Imase, M., et al.: Group-Oriented Communication: Concept and Network Architecture, Computer and Networks, IEEE (2008)
10. Ganshan, W., Yuan, H., Tseng, S.-S., Fuyan, Z.: Knowledge sharing and collaboration system model based on Internet, Systems, Man, and Cybernetics, IEEE (1999)
11. Heissler, J., Kaplan, D., Manoski, S., Wu, J.: Collaborative Planning Over Low Bandwidth Tactical Networks, IEEE (2004)
12. Boodnah, J., Poslad, S.: A Trust Framework for Peer-to-Peer Interaction in Ad Hoc Networks, Computation World, IEEE (2009)
13. Martini, A., Mourao, S., Silva, W.: WhatNow: A System to Enable Videostream in a Mobile Network, Computation World, IEEE (2009)
14. The Symbian Developer Library, `http://developer.symbian.com/.../sdl/` (Accessed: November 26, 2009)
15. Lampton, D., Riley, J., Kaber, D., Sheik-Nainar, M., Endsley, M.: Use of Virtual Environments for Measuring and Training Situational Awareness
16. Brooke Schaab, B., Dressel, J., Sabol, M., Lassiter, A.: Training Collaboration in a Network-assisted Environment. ARI Research Note 2009-05 (2009)
17. Ganguly, A., Agrawal, A., Boykin, P.O., Figueiredo, R.: WOW: Self-Organizing Wide Area Overlay Networks of Virtual Workstations. In: 15th IEEE High Performance Distributed Computing (2006)
18. Hayden, T.W., Ward, C.: Human collaboration tools for net-centric operations, Collaborative Technologies and Systems, IEEE (2005)
19. Behrens, C., Shim, H.: Web Services for Knowledge-Driven Collaborative Modeling, Aerospace Conference Proceedings, IEEE (2004)
20. `http://en.wikipedia.org/wiki/T.120` (accessed 27.12.2009 at 18.47)
21. `http://en.wikipedia.org/wiki/H.323` (accessed 27.12.2009 at 18.01)
22. Raygan, R.E., Green, D.G.: Internet collaboration. TWiki, SoutheastCon, IEEE (2002)
23. Li, J., AlRegib, G.: Optimal Weighted Data Gathering in Multi-Hop Heterogeneous Sensor Networks, San Diego, IEEE (2008)
24. Lewis, G., Morris, G., et al.: SMART: The Service-Oriented Migration and Reuse Technique, Technical report CMU/SEI-2005-TN, Software Engineering Istitute, Carnegie Mellon University (September 2005)
25. Medlow, M.: Extending Service-Oriented Architectures to the Deployed Land Environment. Journal of Battlefield Technology 13(1), 27–33 (2010)
26. Dekker, A.: A Taxonomy of Network Centric Warfare Architectures, `http://www.carlisle.army.mil/DIME/documents/SETEDekker.pdf` (accessed on July 12, 2009)
27. Liu, L., Russell, D., Xu, J., Webster, D., Luo, Z., Venters, C., Davies, J.: Delivering Sustainable Capability on Evolutionary Service-Oriented Architecture, IEEE, Tokyo, pp. 12–19 (2009)

28. De Rango, F., Fazio, P., Marano, S.: Utility Based Predictive Services for Adaptive Wireless Networks with Mobile Hosts, IEEE 58(3), 1415–1428 (2009)
29. Farroha, D., Farroha, B.: SOA as a catalyst to empower the warrior through improved enterprise data access over the GIG, pp. 28–53. IEEE, Vancouver (2009)
30. Saarelainen, T., Jormakka, J.: C4I2-Tools for the Future Battlefield Warriors. In: IARIA/IEEE ICDT 2010, Athens, June 13-19 (2010)
31. Fong, L., Kalantar, M., et al.: Dynamic resource management in an eUtility, network operations and Management Symposium, IEEE, Florence, Italy, pp. 727–740 (2002)
32. Dhanani, S., et al.: A Comparison of Utility Based Information Management Policies in Sensor Networks, Charlottesville, Virginia, pp. 84–89. IEEE, Los Alamitos (2006)
33. Saarelainen, T.: Extending Service Oriented Architecture to Allocation of Battlefield Resources in Tactical Maneuvers. Submitted in Journal of Military Studies

Trust and Reputation Management for Critical Infrastructure Protection

Filipe Caldeira[1,2], Edmundo Monteiro[1], and Paulo Simões[1]

[1] Universidade de Coimbra - CISUC, Coimbra, 3030-290, Portugal
{fmanuel,edmundo,psimoes}@dei.uc.pt
http://www.dei.uc.pt
[2] Polytechnic Institute of Viseu, Viseu, 3504-510, Portugal
http://www.ipv.pt

Abstract. Today's Critical Infrastructures (CI) depend of Information and Communication Technologies (ICT) to deliver their services with the required level of quality and availability. ICT security plays a major role in CI protection and risk prevention for single and also for interconnected CIs were cascading effects might occur because of the interdependencies that exist among different CIs. This paper addresses the problem of ICT security in interconnected CIs. Trust and reputation management using the Policy Based Management paradigm is the proposed solution to be applied at the CI interconnection points for information exchange. The proposed solution is being applied to the Security Mediation Gateway being developed in the European FP7 MICIE project, to allow for information exchange among interconnected CIs.

Keywords: Critical Infrastructures, ICT security, Trust and Reputation Management, Policy Based Management.

1 Introduction

As defined by USA Administration, "Critical infrastructures are those physical and cyber-based systems essential to the minimum operations of the economy and government. These systems are so vital, that their incapacity or destruction would have a debilitating impact on the defense or economic security" [1].

From the above definition, it is clear that CIs (Critical Infrastructures) are one of the most ICT dependent areas of contemporary societies where we should ensure the highest security levels.

Growing interest on this matter is clear from governments initiatives such as the Critical Infrastructure Protection (CIP) Program, started in 1998 by USA Administration, and the European Programme for Critical Infrastructure Protection (EPCIP) in 2006. Also citizens are becoming aware and concerned about this problem due, for example, to a recent television series, "24 -season seven", where the fiction character Jack Bauer fights a terrorist group intending to destroy some USA critical infrastructures. Apart from the fiction involved, this TV series clearly demonstrates how important those infrastructures are and how weak they can be.

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 39–47, 2010.

As stated, European Commission (EC) concerns on CIP result the establishment of the European Programme for Critical Infrastructure Protection (EPCIP) whose main goal is to improve the protection of CIs in the European Union. EPCIP intends to create procedures for European CIs identification; Creation of expert groups; Implementation of the Critical Infrastructure Warning Information Network (CIWIN); CI information sharing frameworks and the identification and analysis of CIs interdependencies. CIWIN main objective is to provide a platform for the exchange of rapid alerts, established by the EC, to help Member States and CI operators to share information on common threats and vulnerabilities.

Recent efforts have been focusing on each CI individually, launching the basis for more secure CIs with enhanced robustness, security and resiliency, introducing, by example, fault-tolerant architectures, redundancy of components and resilient IT systems. One important aspect that still needs to be addressed relates do the interdependency existent among CIs. This interdependency can lead, in a extreme situation, to a global failure started by a single trivial incident in one CI (cascading effect).

Although large efforts have been made in modelling CIs risk analysis, the valuable information gathered from those models are still kept inside each CI and is not shared among interdependent CIs.

In this context the lack of sharing mechanisms for risk information between interconnected CIs was identified. Those mechanisms will allow CI operators to have a real time view on the risk level associated to services on which the modern society depends such as power, water supply, or communication lines. This shared information is also important to increase accuracy of CI risk models introduction external failures risks on those models [3].

The introduction of mechanisms for sharing risk information can, along with more resilient CIs, increase the security level of multiple interdependent CIs. To achieve these service levels, a robust, resilient and inter-dependencies-aware alerting system need to be design and implemented.

This is the main goal of MICIE (Tool for systemic risk analysis and secure mediation of data exchanged across linked CI information infrastructures) FP7ICT project, aiming the design and implementation of a real-time risk level dissemination and alerting system [4].

In this paper we briefly present the MICIE Alerting System, the MICIE Secure Mediation Gateway, and the solutions to specifically incorporate CI interdependencies in the online risk assessment framework. This alerting system is a core component of the MICIE FP7 Project and is in line with the European initiative to establish a Critical Infrastructure Warning Information Network CIWIN [2].

The main contribution of this paper is on the definition of the ICT security mechanisms for information exchange among CIs, namely, the trust and reputation mechanisms to be applied to the Secure Mediation Gateways. The paper proposes the use of the Policy Based Management paradigm to manage the Secure Mediation Gateway.

The rest of this paper is organized as follows. In Section 2 we discuss related work. The key components of the MICIE Alerting System (including the Secure Mediation Gateway) are presented in Section 3. Section 4 presents our aproach to trus and information management. Section 5 discusses advantages and disadvantages of our proposal. Section 6 concludes the paper.

2 Related Work

There are several international projects that have already addressed issues similar the ones targeted by MICIE, but with relevant differences. European projects IRRIIS [5] and CRUTIAL [6] are two relevant projects in this field.

CRUTIAL approach gives particular importance to CIs interdependency modelling to increase CI's resilience to faults and attacks [6]. CRUTIAL main objectives are the development of new architecture and models applicable to heterogeneous CI's; Analysis of applications scenarios in where a small failure could cause a enormous impact in the CI; Research and development of distributed architectures that permit an efficient and reliable control on an electrical network. CRUTIAL expected results will permit gather better knowledge about CIs, permitting the development of more resilient infrastructures [7, 8].

According to [5], IRRIIS intends to develop mechanisms in order to raise the confiability, survivability and resilience of information systems related to CIs. Two main scenarios were defined, representing, respectively, a telecommunications infrastructure and a electricity production and distribution network. For each scenario, the way CI's connect to the exterior by the use of convergent networks like the Internet was analysed [5]. Work was also developed in areas related to on-line risk prediction tools able to incorporate CI interdependencies, studying multiple approaches for data sharing across CIs, interdependency modelling and risk estimators [9].

The IRRIIS project has developed a set of applications named MIT (Middleware Improved Technology). Those applications made possible the communication between CI's using incompatible applications. MIT main objective is to enable a simple, fast and reliable information exchange between CI's, thus reducing response time to incidents that may occur in the CI's by maintaining infrastructure managers well informed about the CI state [5].

Also, in actual research we can find a growing interest in Trust and Reputation areas [10]. Most of the work is focusing on P2P systems [11], wireless sensor networks [12], on-line personal interactions, software agents and in evaluating generic models and formalisms for trust and reputation systems [13].

Both CRUTIAL and IRRIIS projects provided a strong contribution to Critical Infrastructure Protection, but none of them fully addressed the problem of real-time information exchange for real-time CI risk prediction and the security issues associated with the exchange of information, among them adding trust and reputation. Also, for the best of our knowledge this is the first use of policy based management for the control of the trust and Reputation of the exchanged information.

3 MICIE Alerting System

In line with European developments in the Critical Infrastructure Protection field, MICIE project contributes in three main areas: The identification and modelling of interdependencies among CIs, development of risk models and risk prediction tools, and the development of a framework enabling secure and trustfully information sharing among CIs [14].

The main goal of MICIE alerting system will be to provide, in real time, each CI operator with a CI risk level measuring the probability that, in the future, he will loose the capacity of provide some services or receive some service. MICIE overall architecture is presented in Figure 1 [15].

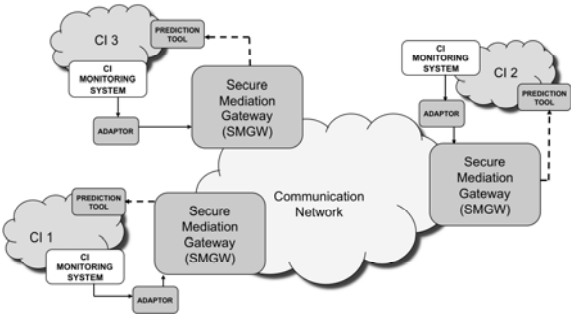


Fig. 1. MICIE overall system architecture [15]

To achieve accurate risk prediction, MICIE system needs information regarding services provided by is own CI and also services provided by interdependent CIs. Following a service-oriented architecture (SoA), we can express interdependencies between CIs also as services. Also internal relations between components or entities in one CI are, in this context, treated as services. Using this approach is possible to introduce the notion of Quality of Service (QoS) expressing the risk of service failure (total or partial, internal or external).

The first step to make risk prediction is to discover all the distributed information that is relevant for the alerting system. Each CI collects, in real-time, information regarding the status of its components. This information is filtered and adapted by an specific adaptor. The adaptor selects proper information and handles format and semantic conversions. Next step is done by the Prediction Tool. This tool make use of risk models to assess the risk level of monitored services. In this entity, CI's own risk level is associated with the risk levels for services received from partner (interdependent) CIs. Each CI using MICIE has at least one Prediction Tool.

Status information as results from the Prediction Tool can be exchanged across partner CIs using a Secure Mediation Gateway (SMGW) allowing CIs to work in a fully cooperative distributed environment for risk prediction.

SMGW main functions can be summarized as: Provision of a secure and trustfully cross-CI communication infrastructure; Collect information about the local CI; Retrieve information about the other interdependent CIs in the system; Send information related to local CI to remote CIs; Composition of CIs critical events and semantic inference; Provide all the collected information to the prediction tool [15].

Each SMGW has a discovery framework implementing dynamic discovery on information available on the local CI and on all interconnected CIs. Besides being used to evaluated risk prediction, this information also provides CI operator a powerful real-time view about identified risks and alerts.

The sensitive nature of exchanged information leave us to take special attention on the security requirements. The SMGW design has to guarantee security requirements,

such as confidentiality, integrity, availability, non repudiation and auditability/traceability. Trust and reputation management are also essential requirements for the information exchange in SMGWs. The proposal of the trust and reputation management mechanisms to be implemented in SMGWs is the main contribution of the present paper.

4 Trust and Reputation Management in SMGWs

Management of the SMGW is the role of the SMGW manager. Developed according to the policy based management paradigm. The SMGW Manager intends to manage all SMGW aspects. SMGW manager also performs monitoring with the help of the Auditing Engine and can also act as Intrusion Prevention and Detection Engine by configuring firewalls in the communication engine. SMGW architecture is presented in Figure 2.

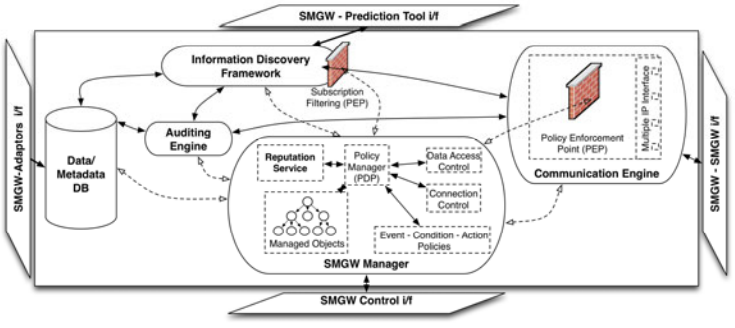


Fig. 2. SMGW Architecture

The proposed approach provides the CI operator with a tool where he can define, in a high level manner, the behaviour he pretends for the system. Traditional approaches are mainly oriented to the management of individual components, not considering the system structure as a whole. In this proposal we use the concept of Policy Decision Point (PDP) -The SMGW Manager -and Policy Enforcement Point (PEP) -The entities that enforce policies, for instance the Communication Engine and the Subscription Filter.

The SMGW Manager handles issues regarding authorization, authentication and accounting controlling both interaction with peer SMGWs and all the internal operation of the SMGW, including also testing, alarming, management of intrusion prevention and detection functions and the management of the Reputation Service.

The CI operator can define policies that will address the relations between local SMGW and foreign SMGWs, including defining how each particular CI can connect and data access policies. The SMGW manager GUI will allow browse existent information and define actions that remote SMGWs can perform (e.g. write and/or read risk information). All data access controls are implemented with a high level of granularity thus maintaining simplicity.

Policies are represented in a formal way using a policy specification language and stored in a policy repository. The SMGW manager interacts with other entities on the SMGW using a dedicated API implemented by a Web Service. The SMGW manager uses Ponder2 toolkit [16] where each SMGW entity is represented using Ponder2 concept of Managed Object. The complete set of SMGW entities will form a Ponder2 SMC (Self Managed Cell). Policy enforcement is based on Ponder2 Authorization Policies and Event Condition Action concepts.

Exploring the features provided with this management approach we improved management aspects introducing the concept of trust and reputation in the context of communication among critical infrastructures.

Although MICIE system can bee seen as a closed system where it is supposed that partners trust each other, it's possible for a partner CI to provide inaccurate information, either maliciously (e.g. if their system is somehow compromised) or due to a faulty components in its monitoring framework.

In this context we have identified the need of a Trust and Reputation Service on each SMGW, able to maintain real time trust information about peering SMGWs. This service will monitor information exchanged between peer SMGWs and partner behaviour in order to manage reputation and to infer a trust level for each one.

There are two main areas where trust and reputation was applied (see Figure 3). First to a trust indicator about the information received from partner CIs. This indicator is evaluated at two levels: Service Level, evaluating each service subscribed to remote CI, reflecting our trust on alerts received on one specific service; and at CI Level, evaluating a trust indicator for each interconnected CI, representing our trust in that particular CI. The Reputation Service is also capable of understand the interdependent CIs behaviour in terms of ICT security. The Interactions between peers are monitored to gather intelligence about the partnership. Thus if one partner CI behaves incorrectly according to defined policies, by example repeating the tentative of retrieve private information, we can see this as a ICT incident and evaluate a trust indicator based on this type of information.

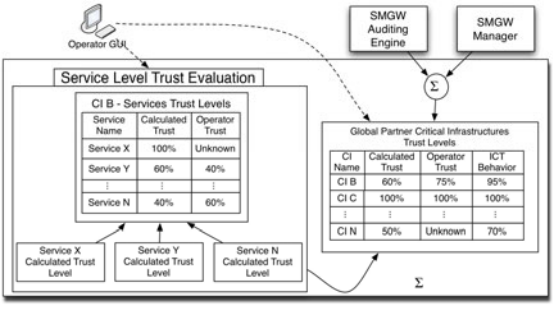


Fig. 3. Reputation Service

Second, a trust indicator is evaluated for each service using information available on the SMGW. As we can keep monitoring services received from peer CIs, we can estimate trust for each service simply evaluating provided risk estimation against actual service behaviour. For example, if a partner CI keeps informing that one service

is going to stop with a probability of 90% and the service never fails we can infer that this particular alert is not credible.

This trust indicator, based on past experience, is evaluated using a statistical approach where operator can define parameters like the penalization he wants to give on detected anomalies between observed service and risk level or the fast or slow aging of past events. This last parameter has special importance to avoid situations where for a long period, received informations is accurate and then starts to be faulty. A simple average will maintain a high trust level while the aging factor will give more attention to the new situation.

During trust evaluation, CI operator can have an active role introducing his own subjective trust regarding specific services or globally about one CI. This indicator introduces aspects not detectable by automatic trust evaluation. By example, CI operator can know that some CI was having faulty equipment during a time period. During this period is likely that our trust indicator decreases. In this case, operator can act, raising his own confidence parameter and consequently don't letting the global trust value decrease in the future.

Gathered trust indicator can then be used to enhance risk prediction models, take decision based on defined policies and also help CI operator to evaluate the relation between his CI and their peers.

5 Discussion

Proposed system intends to enhance contributions expected from MICIE project. MICIE system will provide, in real-time, risk levels measuring the probability that, in the future, a CI will loose the capacity of provide some services or receive some service. This information is based on his own data and on data received from peer CIs. The question we raised is how much we can trust this data? We have confidence in our models but they depended on received information we can't control. So, the answer can be that we trust MICIE risk indicators as much as we trust on information used to evaluate them.

Trust indicators can be incorporated in the Prediction Tool as a mean to improve its accuracy and its resilience to inconsistent information provided by peer CIs making possible, for instance, to give more weight to highly trusted data or ignore data provided by low-trust partner.

Also, system management can become more dynamical with the use of trust and reputation indicators, reacting autonomously when those indicators change. For instance, if our trust regarding the behaviour of one peer decreases below a defined threshold a new policy is triggered and the SMGW can stop accepting connections from that peer.

Prototypes for the SMGW Manager and for the Reputation Service are being developed and we are willing to start simulation tests on this prototypes as special attention needs to be addressed on Trust validation as they will influence our risk models Also, we expect to test our proposal along with MICIE project starting with a simple reference scenario that encompasses a small portion of an electricity distribution network and an interdependent telecommunications network [14]. Planned validation work for the MICIE project will also include more complex scenarios, provided by Israel Electric Corporation and including multiple CIs.

6 Conclusions

This paper reports some first research achievements in FP7 ICT-SEC MICIE project on the development of a real-time risk level dissemination and alerting system. In order to reach MICIE objectives, one of the main key challenge to be addressed is the design and the implementation of a Secure Mediation Gateway (SMGW), namely a new innovative network element able to: (i) discover CI status information, (ii) overcome information heterogeneity and (iii) provide a secure communication of such information among peer CIs. Author's contribution to this enhancement is described in this paper, namely the development of a Policy based Management tool for the SMGW and the incorporation of the concept of Trust and Reputation in the SMGW.

Trust indicators and the use of policies can enhance risk indicators accuracy, help incorporating trust in system management and also help CI operator to evaluate the relation between his CI and their peers.

Authors are implementing prototypes for the presented solutions and will evaluate their work using the demonstrator that MICIE project his willing to develop and test on the field.

Improving MICIE project beyond his initial objectives, described work, represents a step forward in CIs interoperation.

Acknowledgments. Work partially financed by FP7 ICT-SEC MICIE project [4] grant agreement no. 225353, and by the Portuguese Foundation for Science and Technology (SFRH/BD/35772/2007). The authors want to thank all the involved partners for their valuable support to this work.

References

1. Clinton, W.J.: Executive order 13010 -critical infrastructure protection. Federal Register 6I(138), 37347 (1996)
2. Commission, E.: Communication from the commission on a european programme for critical infrastructure protection. COM/2006/0786 final (December 2006)
3. Simões, P., et al.: An alerting system for interdependent critical infrastructures. In: ECIW 2010 -9th European Conference on Information Warfare and Security (2010)
4. Micie: Micie -tool for systemic risk analysis and secure mediation of data exchanged across linked ci information infrastructures. FP7-ICT-SEC-2007.1.7 – 225353 – Annex I – "Descrition of Work" (2008)
5. Irriis project web site (2009), http://www.irriis.org/
6. Crutial project web site (2008)
7. Veríssimo, P., et al.: The crutial architecture for critical information infrastructures. In: de Lemos, R., Di Giandomenico, F., Gacek, C., Muccini, H., Vieira, M. (eds.) Architecting Dependable Systems V. LNCS, vol. 5135, pp. 1–27. Springer, Heidelberg (2008)
8. Dondossola, G., Garrone, F., Szanto, J., Gennaro, F.: A laboratory testbed for the evaluation of cyber attacks to interacting ict infrastructures of power grid operators. In: SmartGrids for Distribution 2008, IET-CIRED (2008)
9. Balducelli, C., Pietro, A.D., Lavalle, L., Vicoli, G.: A middleware improved technology (MIT) to mitigate interdependencies between Critical Infrastructures. Springer, Heidelberg (January 2008)

10. Artz, D., Gil, Y.: A survey of trust in computer science and the semantic web. Web Semantics: Science (January 2007)
11. Chen, S., Zhang, Y., Yang, G.: Trust and reputation algorithms for unstructured p2p networks. In: International Symposium on Computer Network and Multimedia Technology, CNMT 2009, pp. 1–4 (2009)
12. Zahariadis, T., Ladis, E., Leligou, H., Trakadas, P., Tselikis, C., Papadopoulos, K.: Trust models for sensor networks. In: 50th International Symposium on ELMAR 2008, vol. 2, pp. 511–514 (September 2008)
13. Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. Decision Support Systems 43(2), 618–644 (2007)
14. Capodieci, P., et al.: Improving resilience of interdependent critical infrastructures via an on-line alerting system. In: COMPENG 2010 -Complexity in Engineering (2010)
15. Caldeira, F., et al.: Secure mediation gateway architecture enabling the communication among critical infrastructures. In: Future Network and Mobile Summit 2010 Conference (2010)
16. Twidle, K., Dulay, N., Lupu, E., Sloman, M.: Ponder2: A policy system for autonomous pervasive environments. In: Fifth International Conference on Autonomic and Autonomous Systems, ICAS 2009, pp. 330–335 (April 2009)

An Approach to Textual Steganography

Michael Morran and George R.S. Weir*

Department of Computer and Information Sciences,
University of Strathclyde, Glasgow G1 1XH, UK
{michael.morran,george.weir}@cis.strath.ac.uk

Abstract. Textual steganography is a means of concealing an encoded message within text. The appeal in such a system is its potential for hiding the fact that encoding is taking place. The failure to hide the presence of encoding is a form of information leakage and an inherent risk since it suggests that there is important or valuable information in transit. This is considered a major limitation of existing cryptographic techniques as applied to secure information transfer. In this paper, we describe an experimental system that we have developed as a test bed for textual steganography. This system allows us to explore the application of part of speech tagging, word sense disambiguation and synonym replacement as component strategies for textual steganography.

Keywords: textual steganography, information security, textual analysis, natural language processing.

1 Introduction

Information security is vital in today's society [1]. Conventionally, this issue has been addressed through a range of encryption techniques that serve to protect the content of valuable data. Encryption technology has developed to the point that breaking modern encryption schemes is practically infeasible*. The use of encryption however, immediately suggests that certain information must be important. To an adversary, or even the authorities, the presence of encrypted traffic between two individuals may be enough to arouse suspicion. Although the information may be securely encrypted, a third party may be able to intercept the information and either tamper with the message or prevent it ever reaching its destination. This problem will remain as long as encryption is adopted as the sole solution to information security needs.

To illustrate the limitation of an encryption system, and to suggest a more appropriate solution, Simmons presents the "Prisoners' Problem" [2]. The scenario tells of Alice and Bob, two prisoners, who attempt to formulate an escape plan. Any communication between the captives must pass through a warder (Wendy). If any suspicious communication is detected between the prisoners Wendy will not pass on the information. Modern encryption techniques, although secure, result in text which is readily recognisable as cipher text. In this scenario, such techniques would be useless in the presence of Wendy and this indicates the need for information hiding, rather than mere encryption.

* Corresponding author.

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 48–54, 2010.

Various information hiding techniques have been devised over the years. One popular approach is Steganography. A primary goal of steganography is concealing the existence of a message rather than encrypting its content. Considerable work, and research, has been applied to steganography with digital multimedia, especially embedding information in image and audio files (e.g., see [3, 4 & 5]).

Despite the advances in computer technology and the proliferation of multimedia content, the majority of Internet traffic, and much of the information we encounter in our daily lives, is still natural language. In this setting, we sought to explore the potential for steganography based upon natural language.

2 Linguistic Steganography

According to Bergmair [6], linguistic steganography is still very much in its infancy and there is no standard method of performing the functions required to take an intended message and embed it within a meaningful plain text carrier message. A number of diverse approaches have been proposed and central to each is the ability to preserve the meaning of the original text.

2.1 NICETEXT

NICETEXT [7] operates by parsing a cover text and extracting syntactic patterns. This is achieved by using a Part of Speech tagger to produce sentence frames, e.g., [(Noun)(Preposition)(Verb)(Noun)]. A lexicon of words is then collected from the cover text classified by part of speech tags with each word being given a unique binary code. The process of encoding a secret message is then accomplished by selecting a random sentence frame and populating each part of speech tag with an appropriate word from the lexicon based on the secret message bit string.

The resulting output produced by NICETEXT is syntactically correct but suffers from serious drawbacks when considering its semantic and grammatical correctness. Such an approach may be successful in fooling statistical natural language processing methods of steganalysis but any human reader would instantly detect that the output was not semantically correct. As noted by the author himself, the sequence of sentences produced "does not add up to comprehensible speech".

Both Chand & Orgun [8] and Nanhe et al. [9] bring attention to the encryption density required by the NICETEXT system. A system with greater density of encryption, results in less likelihood that the output text will be both grammatical and meaningful. Since a key requirement of any steganographic system is to conceal the presence of a hidden message, the quality of output text is a critical consideration. A later version of NICETEXT, in the form of NICETEXT II, has improved upon the results achieved by NICETEXT but still fails the requirement of generating meaningful text.

2.2 Word Replacement

The word replacement approach exploits the phenomenon of synonyms in natural language. A synonym is a word having the same or nearly the same meaning as another in the language. If a particular word is replaced in a sentence with a synonym then the meaning should be preserved (or closely approximated). As noted by Chand

& Orgun [8] there are limitations in the application of synonyms. Firstly, the number of true synonyms in English is small. Many words have more than one sense and it is rare for two words to be synonymous in all of their senses. Secondly, two synonyms are likely to have very different syntactic distribution and word frequencies throughout a given text or in a particular context. This will be reflected in their conventional association with other words (i.e, in collocations).

Despite these concerns, a word replacement system affords a simple and effective process in linguistic steganography. Bergmair [6] suggests the use of Word Sense Disambiguation to address the challenge of determining appropriate replacement words in order to maintain meaning.

A word replacement system is also able to vary the encryption density. By sparsely distributing the word changes across a large text, the changes are not concentrated to one area of a text. This results in a low encryption density and will therefore appear less conspicuous to a human reader and ensure the text remains as similar to the original text as possible.

3 Our Approach

Our purpose in developing a prototype textual steganography system was to experiment with the use of natural language technologies as a basis for textual substitutions. Our starting assumption was that existing language processing techniques may provide sufficient richness and flexibility to afford the reliability and complexity required of a text encoding facility. In the following, we describe this approach, its degree of success. And what it teaches us about such an enterprise.

In terms of the textual steganography introduction detailed above, we may describe the basis of our system as word replacement combined with word sense disambiguation. This offers the simplest solution to the problem and is an effective means of producing output that is inconspicuous to a human reader.

The first issue to consider is the choice of cover texts. Ideally, there should be broad scope for the implemented system to exploit a variety of cover texts from different genres. A desirable approach lets the user of the system choose their cover text. Since the potential set is drawn from natural language, the choice of cover text becomes almost infinite.

Given a particular cover text, the system then performs word replacement. The proposed system uses a part of speech tagger to tag a given cover text, and then selects particular target words for replacement based on their part of speech and word sense disambiguation. The system targets nouns, verbs, etc and leaves function words, such as determiners, prepositions, etc unchanged. This helps to ensure that the resultant texts retain a grammatical form and meaning.

In order to achieve a ‘sparse’ encoding, each individual letter of the target message is mapped to a replaced word in the carrier text. There is a considerable ratio discrepancy between the carrier text and the text of the intended message. This is the second factor that ensures a sparse encoding. Information, in the form of a ‘key’ is created to indicate which words have been altered in the carrier text. The receiving system requires access to this key in order to decode the message content from the carrier text.

If a secure channel is available, the encoding key can be passed via the secure channel whilst the carrier text is passed via an insecure channel. Any covert listening on the insecure channel may intercept the encoded data but the eavesdropper would be unable to decode the text without knowledge from the secure channel. Furthermore, if the textual steganographic approach is effective, the eavesdropper would not detect the presence of any encoding or encryption. The exact nature of this secure channel is left undefined at this point as there is the freedom to implement this in a variety of different ways.

Such a process underpins our proposed system and allows part of speech tagging and word sense disambiguation to take place on the encoding side. The encoding key, stored on a secure resource, maintains the information required to allow any message to be properly decoded at the recipient side.

Our prototype application consists of a Java-based GUI which orchestrates the encoding process with functional assistance from accompanying Perl scripts. The current version has a combined client application and intermediary server application and integrates with an email facility (IMAP and SMTP). We would expect these components to be separated in a production system.

The main application frame affords the user a number of useful features (Figure 1). User mail accounts are listed in convenient collapsible panels on the left portion of the display. These panels offer a directory tree view of the users chosen IMAP server while the top area of the frame provides a mail listing for the selected directory. This list takes the form of a table, which for user convenience can be sorted via any of the shown attributes.

Messages which have hidden information encoded in the body of the message are clearly indicated in the table. This allows the user to quickly identify any messages which may be a target for decoding. The lower portion of the frame displays the body of any message. A 'Decode' button is provided for eligible messages.

The main frame of the screen provides the user with a toolbar. This toolbar offers a number of common functions, in particular the ability to create a new message. This presents the user with a dialog allowing the user to enter a message, select a cover text, and select the mail server through which to send the mail.

Maximizing configurability was a key requirement for this system in order to ensure greatest flexibility and extend the variety of experimentation. The Settings dialog, as shown in Figure 2, is a vital component of the GUI which ensures users have the ability to modify settings. This dialog permits the user to customize which parts of speech will be considered as targets for word replacement. Some control is also provided on the size of context window (in number of words) used by the word sense disambiguation process. These customiziation dialogues affect the operation of the steganography process and are kept as simple as possible, to afford maximum user friendliness and intuition, while also retaining the ability for advanced users to perform fine-grained modifications. This strategy seeks to provide a happy medium between novice and more experienced users.

The main goals of the final system lay in the ability to hide information in text in such a way that hidden content is undetectable. The analysis and evaluation of the accuracy and performance of the WSD functionality proved to be a significant factor in determining the success of the final system. This analysis was tackled by experimentation and observation. Automated testing was not a practical means of checking that the encoded output was both grammatical and meaningful.

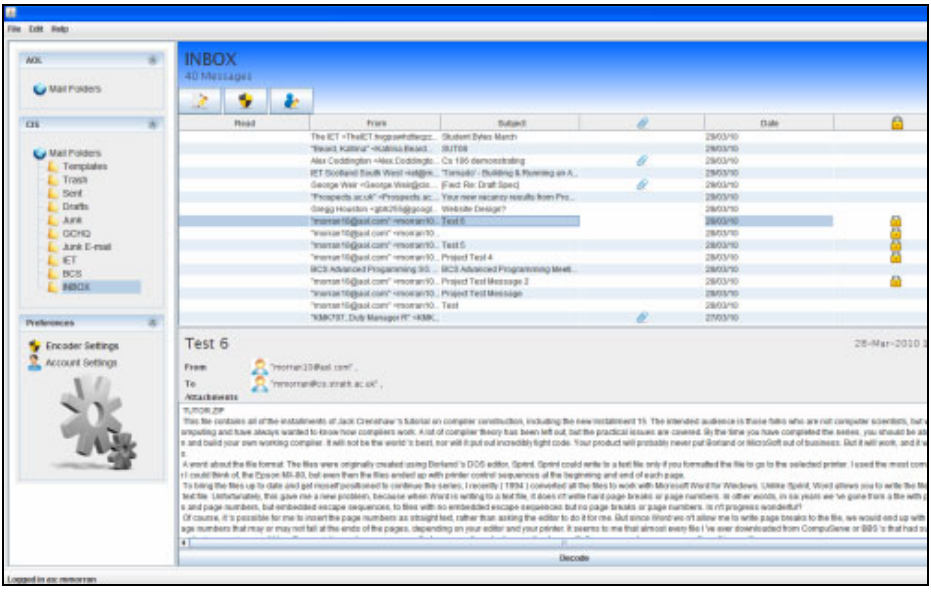


Fig. 1. User interface

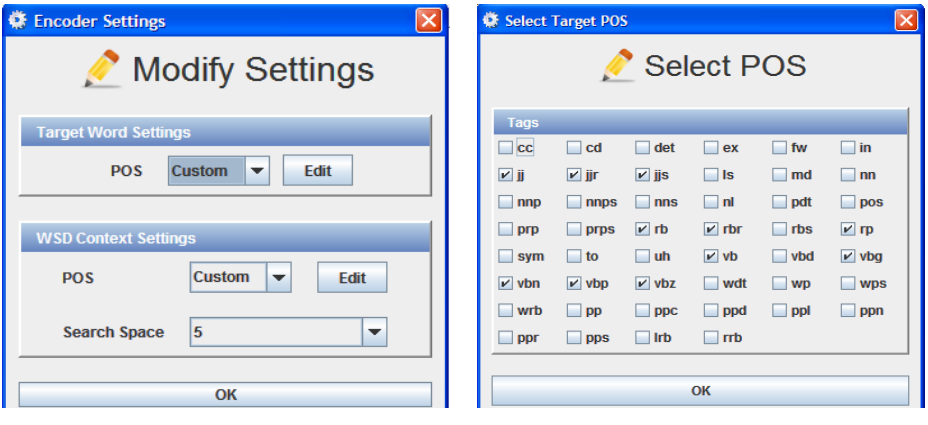


Fig. 2. Option settings

Accuracy in particular is vital in order to produce text which is semantically correct. While many effective part of speech taggers generally claim accuracy in the region of 85-95%, with some suggesting even greater accuracy [9], the general process of automated WSD offers far bleaker statistics. These lower accuracy statistics however relate to determining the sense of every word in a given body of text. Certain classes of words are more prone to ambiguity than others and thereby prove more difficult to map to the correct sense. This distinction between ambiguous and unambiguous words is important and, if unambiguous words are targeted for replacement, the accuracy of the disambiguation function is significantly increased.

A major feature of our prototype is to allow configurability in the word replacement system. Much of this configurability relates to the process of WSD and lies in the ability to modify targeted word groups, search space scope and choice of context words. To attempt to produce an optimal strategy, testing was carried out on this set of configurable variables in order to determine the effects of differing configurations.

4 Conclusions

A series of tests were carried out to determine the impact of the language processing components in the prototype steganography system. Bearing in mind that we were able to achieve full recovery of any encoded message, the primary consideration under evaluation was the viability of hiding the encoding. In this regard, the accuracy of the WSD function proved to be a major issue. Review of the resultant replacement texts led to the conclusion that default settings would generate acceptably meaningful substitutions in 50%-60% of cases. The remaining cases were likely to be seen as 'odd' by the average reader.

While these results indicate that this automated word replacement system is not able to guarantee meaningful output in all instances, they show that by allowing configurability, existing techniques can be used in conjunction with optimal replacement strategies to produce a viable and effective solution. The results indicate that tailoring a particular strategy to the context of a chosen cover text can enhance the scope for optimal replacement.

Clearly, accurate creation of grammatical and meaningful output is a desirable factor in the perceived success of the final system, yet different human readers were found to have different interpretations of which substitutions result in errors. This was proven by three test cases that asked a number of readers to verify the text produced after substitution had taken place. The results were surprising. In fact, in most cases readers failed to spot the majority of substitutions despite being told in advance that

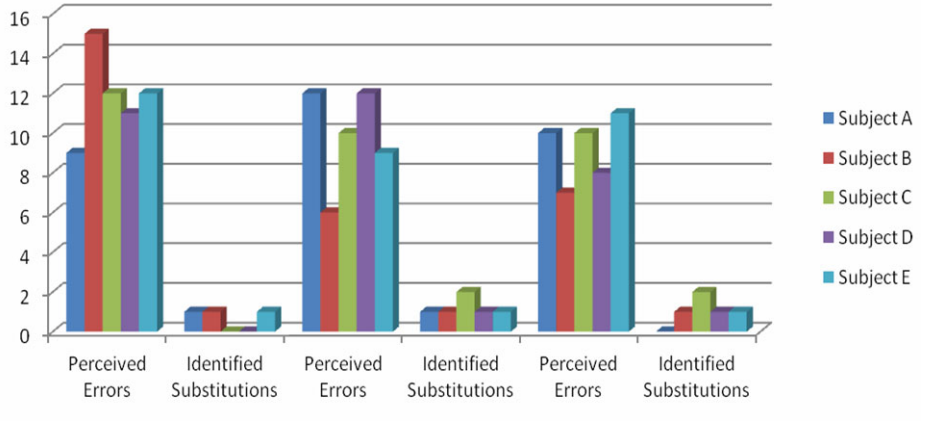


Fig. 3. Results of user testing

there would be substituted words. In all cases, more perceived errors were identified than actually existed (Figure 3). In test case 1, there were seven word substitutions. Test case 2 also had seven word replacements, while test case 3 had ten word substitutions. These results illustrate the diverse judgments made by different individuals and exemplifies some of the difficulties arising from lexical ambiguity.

This 'human factor' does however provide promising results in the context of our system and reflects well on the use of sparse encryption density. With sparse encryption density, many readers simply do not notice slight errors within large blocks of text. Analysis of the results suggests that readers simply interpret any semantic ambiguities with a meaning which they are expecting, given the particular context of the text. Perhaps this indicates that a small degree of 'semantic noise' will not be noticed, or be tolerated and ignored. Coupled to this, as automated disambiguation abilities improve, so too will the ability of such linguistic steganography systems to produce meaningful and inconspicuous text.

References

1. Whitman, M.E., Mattord, H.J.: Principles of Information Security, 3rd edn. Broadman & Holman Publishers, Tenessee (2007)
2. Simmons, G.J.: The Prisoners Problem and the Subliminal Channel. In: Advances in Cryptology: Proceedings of Crypto 1983, pp. 51–67. Plenum Press, New York (1983)
3. Chandramouli, R., Memon, N.: Analysis of LSB based Image Steganography Techniques. In: Proceedings of ICIP (2001)
4. Johnson, N.F., Jajodia, S.: Exploring Steganography: Seeing the Unseen. IEEE Computer 31, 26–34 (1998)
5. Hunt, K.: A Java Framework for Experimentation with Steganography. In: Proceedings of the 36th SIGCSE technical symposium on Computer science education, SESSION: Programming with images, pp. 282–286 (2005)
6. Bergmair, R.: A Comprehensive Bibliography of Linguistic Steganography. In: Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents (2007)
7. Chapman, M., Davida, G.: Hiding the Hidden: A Software System for Concealing Ciphertext as Innocuous Text. In: Han, Y., Quing, S. (eds.) ICICS 1997. LNCS, vol. 1334, pp. 333–345. Springer, Heidelberg (1997)
8. Chand, V., Orgun, C.O.: Exploiting linguistic features in lexical steganography: design and proof-of-concept implementation. In: Proceedings of the 39th Annual Hawaii International Conference on System Sciences, vol. 6, p. 126b. IEEE, Los Alamitos (2006)
9. Nanhe, A.M., Kunjir, M.P., Sakdeo, S.V.: Improved Synonym Approach to Linguistic Steganography Design and Proof-of-Concept Implementation (2008)
10. Cutting, D., Kupiec, J., Pedersen, J., Sibun, P.: A Practical Part of Speech Tagger, Applied Natural Language Conferences. In: Proceedings of the Third Conference on Applied Natural Language Processing, pp. 133–140 (1992)

Cybercrime Victimisations/Criminalisation and Punishment

Ameer Al-Nemrat, Hamid Jahankhani, and David S. Preston

University of East London
School of Computing, IT and Engineering, UK
ameer@uel.ac.uk, hamid.jahankhani@uel.ac.uk

Abstract. With the increased of use of the internet as a means of sharing information, the need to protect and preserve the confidentiality and integrity of data is ever more evident. The digital age provides not only established criminals with new ways of committing, but also has empowered previously non deviant individuals, into new cyber criminal behaviour. Many individuals are unaware of online threats and many fail to take advantage of precautionary measures to protect themselves from risks when they are online. Therefore, individuals consistently underestimate their risk of becoming victims or underestimate the punishment that may face if they are engaged on online deviant behaviour. This ongoing research has found that there is a relationship between individual's perception of cybercrime law and cybercrime victimisation and/or criminalisation.

1 Introduction

The onslaught of cyber crime has been difficult to measure with the ever increasing global capacity of the internet and free access to it. The anonymity of offenders, hiding behind International barriers, utilising weaknesses in cross boarder jurisdictions and inadequate 'mutual assistance' agreements between Countries around the World, all play an important part in assisting would be offenders.

Victims in the United Kingdom and around the World often find themselves with no means of recourse at a local level. Police authorities have little or no means of investigating e-crime and budgetary controls ensures that only the most serious offences (usually Terrorist related or of a sexual connotation) will ever be investigated.

The impact and influence of the Internet over the past 12 years has been immense. During that time, access to the Internet has grown enormously. In 1996, 3.4 million UK adults were online; by 2006 this had expanded to 28.5 million, "UK Cybercrime Report", (Trend Micro, 2008).

The prevalence of the Internet, its sheer enormity and exponential growth has revolutionised global communications at an unprecedented scale. The manner in which we conduct our business and private affairs have changed drastically with the advent of Technological advances – particularly with the open accessibility to e-mail.

The information technology age provides not only established criminals with new ways of committing crime, but also has empowers previously non deviant individuals, drawing them into new criminal behaviour (Williams, 2005). According to Trend Micro, Figure 1, in 2008 there was an increase in overall Web threat activity - these

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 55–62, 2010.

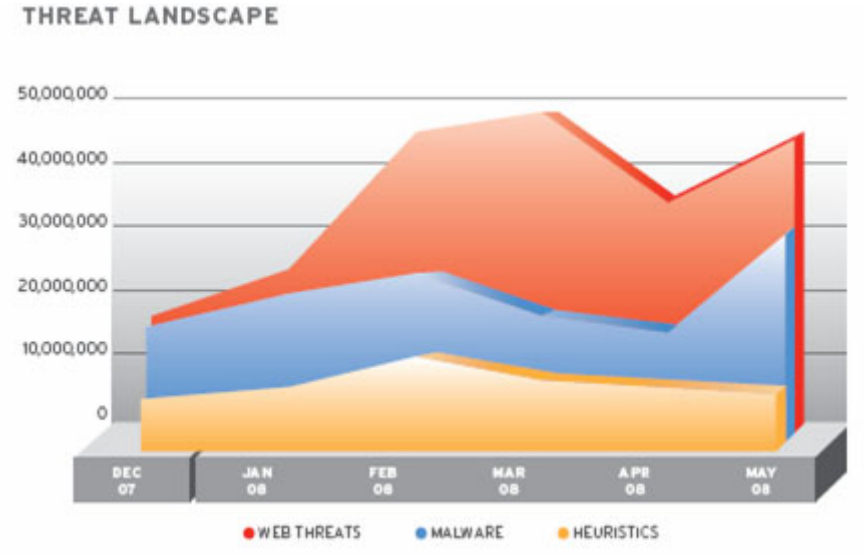


Fig. 1. Web threat landscape 2008, (Trend Micro, 2008)

new forms of Crimes has posed a new challenge for computer security specialists, legal professionals, information managers and law enforcement agencies at all levels on how to prevent, investigate and prosecute.

The term 'Cybercrime' has been used to describe a number of different concepts of varying levels of specificity. At its absolute broadest, the term has occasionally been used to refer to any type of illegal activity that result's in a pecuniary loss. This would include violent crimes against the person or property such as armed robbery vandalism, or Blackmail.

At its next broadest, the term has often been used to refer only to non-violent crimes that result in a pecuniary loss (Pease 2001). This would include crimes where a financial loss was an unintended consequence of the perpetrator's actions, or where there was no intent by the perpetrator to realize a financial gain for himself or a related party (e.g. When a perpetrator hacks into a bank's computer and either accidentally or intentionally deletes an unrelated depositor's account records.).

Although the term 'cybercrime' is now in everyday use, the first problem encountered in *measuring* cybercrime is that there is no commonly-agreed definition of the term.

Despite the fact that the word "Cybercrime" has entered into common usage, many people would find it hard to define the term precisely. Furthermore, there is no catch-all term for the tools and software which are used in the commission of certain online crimes (Ivanova, 2006).

Cybercrime has become a common term and its usage tends to generalise just about any illegal activity within the Internet environment (Ainsworth 2001). Despite an apparent acceptance of and familiarity with the term, there exist dramatically varied views of what Cybercrime *is*. This lack of definitional clarity is problematic as it impacts every facet of e-crime investigation and the reporting process.Some of the definitions of cybercrime that do loosely exist include (Rogers et al 2006).

Cybercrime is described as criminal activity in which the computer or network is an essential part of the crime, this term is also used to include traditional crimes in which computers or networks are used to enable the illicit activity (Kabay and Bosworth 2005).

Other examples of cybercrime refer to where the computer or network is used as a tool of the criminal activity includes spamming and criminal copyright crimes, particularly those facilitated through peer-to-peer networks, or where the computer or network is a target of criminal activity include unauthorized access (i.e., defeating access controls), malicious code, and denial-of-service attacks.

Cybercrime could also encompass where the computer or network is a place of criminal activity which involves the theft of a service (in particular, telecom fraud) and certain financial frauds.

Finally, examples of traditional crimes facilitated through the use of computers or networks include Nigerian (419) Frauds or other gullibility or social engineering frauds (e.g., hacking "phishing", identity theft, child pornography, online gambling, securities fraud, etc.).

Cybercrime in the context of national security may involve hacktivism (online activity intended to influence policy), traditional espionage, or information warfare and related activities (Hunt 1994).

One of the recent researches showed that a new cybercrime is being perpetrated every 10 seconds in Britain, [GSO, 2009 a]. During 2006 the computer crooks were able to strike 3.24 million times. Some crimes performed on-line even surpassed their equivalents in real world. In addition, experts believe that about 90% of cybercrimes stay unreported (GSO, 2009 b).

According to a study performed by McGuire, (NCL 2007), a specialist in psychology, the University of San Francisco, the majority of teenagers who hack computer systems are doing it for more fun than aiming to cause harm. McGuire, [5], also, reports that, "*often parents cannot understand the motivation of the teenage hackers"*. McGuire, conducted an anonymous experiment in the area of San Diego by questioning more than 4,800 students.

The results of the McGuire Survey, (NCL 2007), were presented at the American Psychological Association conference as follows;

- *18% of all youngsters confessed of entering and using the information stored on other personal computer or website;*
- *13% of all the participants mentioned they performed changes in computer systems or computer files.*
- *The study revealed that only 1 out of 10 hackers were interested in causing certain harm or earns money. Most teenagers performed illegal computer actions of curiosity, to experience excitement.*
- *38% of teenagers were involved in software piracy;*

Many previous researchers attempted to understand the cyber criminal's behaviour. There is a lack of insight into the relationship between cybercrime victims who used public computers (Cybercafé) to go online and their perception of punishment.

2 Methodology

2.1 Participants

The analysis utilises data from a face to face survey administered to 232 cyber café users of four different cyber café in London on December 2008. (61.2%) of the respondents were male and (38.8 %) were female (table 1). Furthermore, participants are representative of the larger population of interest (i.e., individuals over 26 years of age who suppose to be knowledgeable of risky behaviours regarding computers and other technological devices (Skiner and Fream, 1997, Higgins, 2005, Holt, 2009). And less than 25 years who were born in the internet age which starts 1982.

The questionnaire was particularly interested about 6 different common trends of cybercrime categories which are; Internet Fraud, Identity theft, Hacking, Online stalking, Viruses, and Phishing.

Table 1. Provides a general overview of respondent demographic

Variables		Frequency	Valid %
Gender			
Male		142	61.2
	Female	90	38.8
Age			
Under 25 years		132	55.5
	Over 26 years	106	44.5
Education			
	Degree	131	55.0
	None	107	45.0
IT Skills			
	Beginner	31	12.9
	Intermediate	119	49.6
	Advanced	72	30.0
	Expert	18	7.5

Data mining was used in the analysis of the study data. Data mining which defined as the exploration of sets of data (usually large ones) in order to discover their features, pattern, and regularities that can be useful in solution of practical and Theoretical problems (Justickis 2009, P657) is used. Information produced by data mining techniques can be represented in many different ways.

Decision trees are one of the most popular methods for learning and reasoning from feature-based examples. "A decision tree is a flow-chart-like tree structure where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and leaf nodes represent classes or class distribution"(Han and Kamber, 2001). They have been deeply studied in both areas of pattern recognition and of machine learning. A pattern is classified by starting at the root node of the tree, testing the attribute specified by this node, then moving down the tree branch corresponding to the value of the attribute in the given example. This process is then repeated for the subtree rooted at the new node until a leaf is encountered, at which time the pattern is asserted to belong to the class named by that leaf.

In this research, in order to better understanding the relationship betweenthe Q12(Have you ever been the victim of one of these types of cybercrime?

- Internet Fraud
- Identity Theft
- Hacking
- Online Stalking
- Viruses
- phishing

cybercrime types) and the rest of the questionnaire' observed variables classification-tree induction was used. Given the interpretability of classification trees, this approach was applied to the data using the CARTapproach to tree induction (Breiman et al., 1984).

Software used for tree induction was the *rpart* library from the R statistical package, which is based on Breiman, Friedman et al. (1984). The default settings of the *rpart* library were used, except that 50-fold cross-validation was used instead of 10-fold cross-validation.

2.2 Hypothesis

This research proposes that cybercafé users are subject to experience a proportionally higher level of victimisation or criminalisation than they would experience in any other places. Therefore, online users who used cybercafés and have risky online behaviours, and underestimate law respond to cybercrime are more likely to be victimised or engaged to some of illegal activities.

3 Results and Discussions

The data understudy produced 4 trees. However, for the purpose of this paper, only one tree is shown in this article (Figure 2).

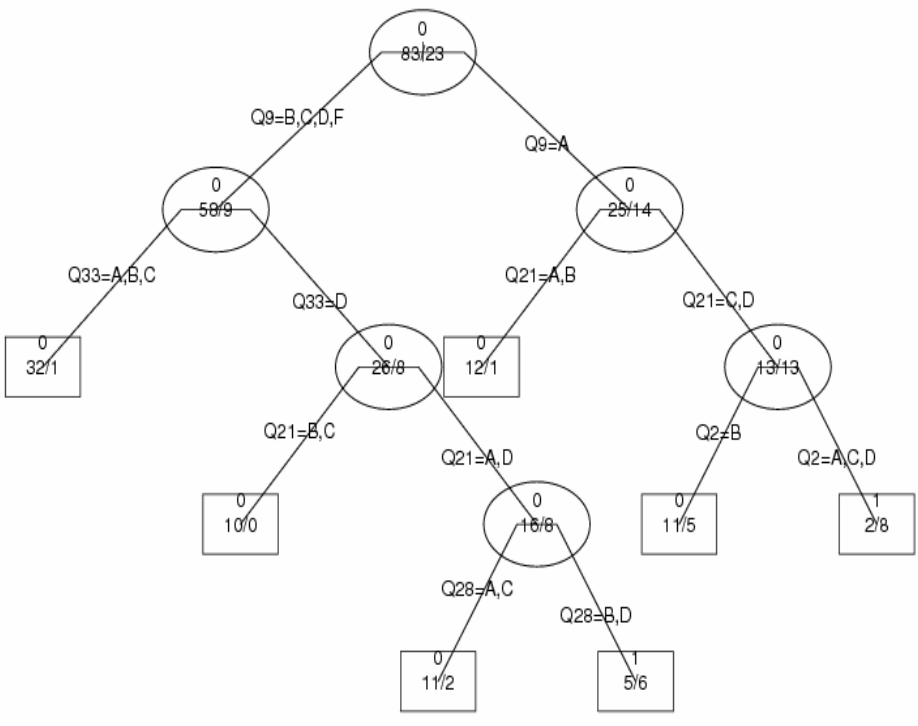


Fig. 2. Classification tree for Internet fraud with respect to UK data. Each node shows the majority class (i.e., 1 if the cybercrime category of interest is present; 0 if absent) and the frequency distribution of the classes at the node :< frequency of class 0> / < frequency of class 1>. Q9 refers to Question 9.

Tree 1 UK user most probably victim of Internet fraud

if {their most frequent use of Internet is not for their job (Q9)

and they agree that the justice system does not treat computer crime as seriously as street crime (Q33)

and they either agree or strongly disagree that they are more comfortable using an Internet café when visiting unknown websites than when using their own computer (Q21)

and they believe that punishment should be at least severe if they are caught with destructive malware (Q28) }

or

if {their most frequent use of Internet is for their job (Q9)

and they at least agree that they are more comfortable using an Internet café when visiting unknown websites than when using their own computer (Q21)

and they are either under 25 or over 26 years old (Q2) }.

Based on the result obtained above it is believed that the most effective measures to reduce cybercrime victimisation has to starts from two sources; end users awareness and law that should govern online behaviour. There is a large body of evidence, (Finne 1987, 2000, Boss 2007) suggesting that the decision to take precaution stems from individual perception of risk. Furthermore, According to Eurim 2003, personal awareness and public education of cybercrime methods and measures remains the first line of defence against cybercrime. In addition, the rapid change of technology and the weaknesses in the legal systems to cope with these changes increased the importance of individual awareness to protect them of cybercrime. Although a new generation is growing up that is computer literate and technology aware, even they are not being taught about the responsibilities that associated with the use of this technology (Eurim 2003).

4 Conclusion

This research has found that cybercafés users are subject to experience a proportionally higher level of victimisation or criminalisation than they would experience in any other places. Therefore, online users who frequently use cybercafés and have risky online behaviours, who lack of awareness of risk to take precaution measures, are more likely to be victimized or engaged to some of illegal activities**.**This finding was assessed using data mining technique (Classification Trees).

References

Ainsworth, P.B.: Offender Crime Profiling and Crime Analysis. Willan Publishing, USA and Canada (2001)

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and Regression Trees. Chapman & Hall, New York (1984)

Boss, et al.: Familiarity Breeds Content: How Fear of Cybercrime Influences Individual Precaution-Taking Behaviour. In: IFIP TC 8 International Workshop on Information Systems Security Research (2009)

EURIM: Working Paper 1: Reporting Methods and Structures; IPPR E-Crime Study. Partnership Policing for the Information Society (2003),
http://www.eurim.org/consult/ecrime/dec03/ECS_WP1_web_031209.htm

GSO. GetSafeOnline (2009a),
http://www.getsafeonline.org/nqcontent.cfm?a_id=1143

GSO. GetSafeOnline (2009b),
http://www.getsafeonline.org/nqcontent.cfm?a_id=1499

Holt, J.T., Bossler, A.M.: Examining the applicability of Lifestyle- routine activities theory for cybercrime victimisation. Deviant Behaviour 30, 1–25 (2009)

Hunt, D.: Preventing Criminal Victimisation: The case for an Intersectoral Response to Victimisationa South Australian Perspective (1994),
http://www.aic.gov.au/publications/proceedings/27/hunt.pdf

Ivanova, P.: cybercrime & cyber security. Information & Security, an international Journal 18(1) (2006)

Justickis, V.: Criminal data Mining. In: Jahankhani, H., et al. (eds.) Handbook of electronic security and digital forensics. World Scientific, London (2010)

Kabay, M.E., Bosworth, S.: Computer security Handbook. John Wiley & Sons, Inc., New York (2005)

NCL's Fraud Centre Survey (2007),
http://www.fraud.org/internet/intstat.htm

Pease, K.: Crime Future and foresight. In: Wall, D.S. (ed.) Crime and the Internet. Routledge, London

Rogers, M., et al.: Self-reported computer criminal behaviour: a psychological analysis. Digital Investigation Journal (2006), http://www.elsevier.com/locate/diin

Trend Micro: Securing your web (2008),
http://us.trendmicro.com/imperia/md/content/us/pdf/threats/securitylibrary/1h_2008_threat_report_final.pdf

Williams, M.: Cybercrime. In: Miller, M. (ed.) Encyclopaedia of criminology. Routledge, London (2005)

Williams, K.S.: Using title control balance theory to understand computer crime and deviance. International review of Law Computers & Technology 22(1), 145–155 (2008)

Baby-Crying Acceptance

Tiago Martins[1] and Sérgio Tenreiro de Magalhães[2]

[1] Universidade Católica Portuguesa, Braga, Portugal
tiagorsmartins@hotmail.com
[2] Universidade Católica Portuguesa, Braga, Portugal
stmagalhaes@braga.ucp.pt

Abstract. The baby's crying is his most important mean of communication. The crying monitoring performed by devices that have been developed doesn't ensure the complete safety of the child. It is necessary to join, to these technological resources, means of communicating the results to the responsible, which would involve the digital processing of information available from crying. The survey carried out, enabled to understand the level of adoption, in the continental territory of Portugal, of a technology that will be able to do such a digital processing. It was used the TAM as the theoretical referential. The statistical analysis showed that there is a good probability of acceptance of such a system.

Keywords: baby, crying, safety, technology, acceptance.

1 Introduction

The human being has always expressed a great need to communicate. The main instrument of communication is the language, whatever it is.

So that the children, in their non-verbal stage, can be correctly understood, it is essential to understand the language they use. The crying, the laughter and the lalation are means of communication with other people.

Among the child ways for communication, the most troubling is the crying. It may seem that the baby cries just for the pleasure of crying, but there are several reasons that lead him to cry: hunger, thirst, discomfort, need for sleep, pain, suffering [1].

The most common causes for crying are infections and colic. Other less common appear associated with foreign body ingestion, neurological diseases, reactions to medicines and vaccines, leukaemia, heart or metabolic problems, gastroenteritis, stomatitis, sepsis or other serious bacterial infections [2], myasthenia gravis, congenital myasthenic syndrome [3], changes in the larynx, in the neurological system and genetic syndromes [4], West syndrome - a severe form of epilepsy [5], hypothyroidism [6], meningitis [7],. Crying was also observed in children who subsequently were Sudden Infant Death Syndrome victims [8].

Many of the diseases that manifest themselves through the crying require a medical emergency, a quick diagnosis, so these baby demonstrations should be understood as an warning for an early intervention that may avoid serious sequels and even the death.

In order to preserve the baby safety, have been being developed devices that alert parents to potential dangers of their children, namely walkie talkie (safety monitoring)

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 63–70, 2010.

[9], the Baby Alarm (forgetting the baby in the car) [10], Baby Watch (monitoring of breathing and movement), surveillance cameras (monitoring of various spaces) [11] and voice recognition applications[12].

However, the crying monitoring capacity is not enough, since the responsible people for babies are often separated of them. Thus, it would be necessary to join to the development of technological resources, some ways to communicate, to the responsible for babies, the results, which would imply the digital processing information available by the crying. Intrinsically, this is a call for giving to the current means of communication available to parents to monitor their children, new capabilities, new symbols and rules that define the need for warnings.

The number of highly sophisticated technologies has been growing too quickly. However, the technical characteristics of an innovation, even promising a very good performance, are not by themselves guarantee that it will be successful. The effective use of such a technology by the users is undoubtedly a success factor in its implementation [13].

It is difficult to estimate the future use of a new technology during its development, since it does not exist. However, it is possible to evaluate the individual intention to use it, and there is evidence that the intention to manifest a particular behavior predicts the future behavior [14].

With regard to the babies crying, it was tried to understand the level of adoption, in the continental territory of Portugal, of a technology - fictitious name BabyCrying - which can detect the baby crying, as well as the reason he cries, with features that permit to send warnings in real-time (eg SMS) and remote access to information on statistical reports (how many times the baby cried, why, what time it happened, how long the crisis of crying).

2 Theoretical Referential

Over the past decades, many innovations adoption models have been developed to determine the responsible factors for the success of new technologies adoption. For this investigation was adopted as a theoretical referential the Technology Acceptance Model (TAM), under which two relevant beliefs, the perceived usefulness and perceived ease of use, are the effects of external variables mediators in attitude and in the actual use of the system. The model can be expressed using the following four equations [15]:

$$EOU = \sum_{i=1}^{n} \beta i \, Xi + \varepsilon \; ; \tag{1}$$

$$PU = \sum_{i=1}^{n} \beta i \, Xi + \beta n+1 \, EOU + \varepsilon \; ; \tag{2}$$

$$ATT = \beta 1 \; EOU + \beta 2 \, PU + \varepsilon \; ; \tag{3}$$

$$USE = \beta 1 \; ATT + \; \varepsilon . \tag{4}$$

where
Xi = design feature i, i=1,n
EOU = perceived ease of use
PU = perceived usefulness
ATT = attitude toward using
USE = anticipate system use
βi = standardized partial regression coefficient
ε = random error term.

Several reasons contributed to this choice: it is a specific model for information systems, it has a strong theoretical basis, in addition to an extensive empirical support through validations, applications and replication, because it has already been tested in different samples and in different situations, proving to be valid and reliable and provide strength to map the impact of external factors on the internal ones, in the individuals, in relation to accept or not accept the information technology.

3 Methodology

The choice of methodology to conduct the investigation process to which this research is concerned fell in the survey, made through inquiry based in questionnaire.

Having been chosen an area random sampling, five regions were considered in the continental territory of Portugal (North, Centre, Lisbon, Alentejo and Algarve), in which 1,000 questionnaire copies were distributed, according to the percentages of population in each of them. From these copies, only 600, selected at random among the valid ones, would be analyzed, since this was the value found for a sample size representative of the continental territory of Portugal population, for a 95% confidence level in the results and a 4% sampling error.

The questionnaire consists of two parts: the first one intends to collect data to characterize the respondent and the second includes the variables defined to achieve the proposed objectives, according to TAM. These variables were constructed based on instruments validated in previous studies, particularly in Davis questionnaire. And like in this one, a Likert scale of 7 points was used. This is a psychometric response scale, standardized, commonly used in questionnaires. If the filter question answer was negative, there was an instruction in order to stop filling the questionnaire [16].

4 Statistical Analysis

To obtain all the values as well as to carry out the analysis that refers to this research, it was used the statistical software SPSS (Statistical Packet for Social Sciences).

4.1 Descriptive Analysis

According to the answers given to the questions from the first part of the questionnaire, were made the distributions by Sex, Age, Region and Qualification, of the elements selected at random to form the sample.

4.2 Principal Components Analysis

Although the data collection instrument has been the target of several validations, the Cronbach's coefficient alpha confirmed, once again, its reliability, since for this coefficient and for the four main variables (Perceived Ease of Use, Perceived Usefulness, Attitude Toward Using and Anticipate System Use), were found the values listed in the table 1 and because they exceed 0.8, they show a good or very good internal consistency of the items used in the evaluation of each one of the different dependent variables [17].

Table 1. Reliability Statistics

Variable	Label	N of Items	Cronbach's Alpha
Perceived Ease of Use	EOU	7	0.914
Perceived Usefulness	PU	6	0.947
Attitude Toward Using	ATT	3	0.870
Antecipate System Use	USE	4	0.822

In order to get the values of latent variables Perceived Ease of Use, Perceived Usefulness, Attitude Toward Using and Antecipate System Use, it was used the Principal Components Analysis, which are intended to identify linear combinations of the observed variables that capture the maximum variability of the latter ones, with the lowest possible number of components. However, it was necessary to determine, before, if the data characteristics were suitable to the practice of this type of analysis. So were performed the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) and Bartlett's Test of Sphericity.

The values obtained for the Kaiser-Meyer-Olkin Measure of Sampling Adequacy and Bartlett's Test of Sphericity and for each one of the four main variables are found in the table 2.

Table 2. KMO and Bartlett's Test values

Variable		EOU	PU	ATT	USE
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.837	0.918	0.733	0.646
Bartlett's Test of Sphericity	Approx. Chi-Square	3579.568	3667.681	1466.727	1133.073
	df	21	15	6	3
	Sig.	0.000	0.000	0.000	0.000

Comparing the simple correlations with the partial correlations, the KMO test measures the adequacy of variables to the model, having returned, in the case of this research, values exceeding 0.5, indicating that the factor solution can be applied to the data. Bartlett's Test of Sphericity tests the null hypothesis that the correlation matrix is an identity matrix. As the level of significance obtained with this test is 0.000 (less

than 0.05), the null hypothesis is rejected, meaning that the correlation matrix is not the identity, and therefore there is a correlation among the variables.

Considering the results of these two tests, it's possible to infer that the different sets of independent variables fit very well to the treatment through the Principal Components Analysis.

What is intended is the determination of each one of the independent variables weights in the respective dependent variable. So, were analyzed separately, the different sets of independent variables defined in the questionnaire, which was based on another with successive validations. Satisfying the criterion that the eigenvalue is greater than 1, from each analyzed set was only extracted one principal component which the name is the one of the corresponding dependent variable.

Perceived Ease of Use. In the case of the dependent variable Perceived Ease of Use, with the values found for the total of the explained variance by the different factors, for the eigenvalues, and for the independent variables weights in this dependent variable, it can be concluded that the only extracted factor explains 67.656% of the total variance of the original variables and that EOU can be written as the following linear combination:

$$\begin{aligned} EOU = {} & 0.849*B7 + 0.856*B8 + 0.871*B9 + 0.759*B10 + 0.799*B111 + \\ & + 0.770*B112 + 0.847*B12. \end{aligned} \tag{5}$$

This expression enabled to calculate the EOU values through its observed principal components values.

Perceived Usefulness. As in the previous case, for the Perceived Usefulness was extracted only one factor which explains 80.063% of the total variance of the original variables and with the obtained weights it's possible to write the following linear combination, which enables to calculate the variable PU values through its observed principal components values:

$$PU=0.926*C13+0.927*C14+0.793*C15+0.909*C16+0.915*C17+0.892*C18. \tag{6}$$

Attitude Toward Using. About Attitude Toward Using it´s possible to infer that, once more, was extracted a single factor that explains 72.713% of the total variance of original variables and the obtained weights enable to write the following linear combination that gives the ATT values.

$$ATT = 0.856*A61 + 0.874*A62 + 0.789*A63 + 0.888*A64. \tag{7}$$

Antecipate System Use. Finally, for the variable Antecipate System Use the single extracted factor explains 77.818% of the total variance of the original variables and, with the obtained weights, it's possible to write the following linear combination, which will enable to calculate the USE values:

$$USE = 0.940*D19 + 0.769*D20 + 0.927*D21. \tag{8}$$

4.3 Linear Regression

Obtained the values of the dependent variables, was carried out a succession of multiple linear regressions, in order to arrive to a model for predicting the dependent variable Antecipate System Use. However, this model for analysis can be only used for estimation and inference of functional relationships among the dependent and the independent variables if the residuals are random, independent and follow a normal distribution with mean zero and constant variance across observations (homoscedasticity) and the independent variables are independent from each other, that is, they are not correlated.

Assumptions Validation. For each one of the dependent variables, the residuals randomness and the homoscedasticity assumptions were validated graphically with scatter plots. The residuals distribution normality was validated with normal probability plots. The residuals independence was validated with the Durbin-Watson's *d* statistic, that tests the null hypothesis that the residuals are uncorrelated against the alternative hypothesis that there is significant serial correlation among the residuals (as a rule of thumb, values of $1.5 < d < 2.5$ show that there is no autocorrelation in the data). In addition it was checked for outliers (observations which the absolute value of the Studentized Deleted Residual exceeds 1.96) and it was made its elimination, since linear regression is sensitive to outlier effects. The independent variables multicollinearity was validated with the Variance Inflation Factor (VIF) (values of this factor exceeding 5 indicate the presence of multicollinearity in the independent variables).

After the assumptions validation, it was carried out several linear regressions to model the relationship among each one of the dependent variables and the respective independent variables.

5 Analysis and Discussion of Results

After the assumptions validation, it was carried out a linear regression that led to the results that are in the tables below:

Dependent Variable EOU. In regard to the variable Perceived Ease of Use, all the regression coefficients are significant (p_value <0.05) and the fitted model explains 73.9% of the variability of this variable. With the Standardized Beta Coefficients it was possible to write the following structural equation for this variable:

$$EOU = 0.112*E22 + 0.442*E23 + 0.361*E24 + \varepsilon. \quad (9)$$

Dependent Variable PU. An initial linear regression analysis performed on the dependent variable Perceived Usefulness, entering all independent variables, revealed some very high VIF values, associated with p_values exceeding 0.05, which may be an indication for multicollinearity to be present. Therefore, it was carried out a new regression using the stepwise method, which eliminated the non-significant variables for the model.

The new VIF values validated the independent variables entered in the model independence assumption. All the coefficients of these variables are significant and the

adjusted model explains 79.1% of the variability of this variable. Using the Standardized Coefficients Beta, results the following structural equation for PU:

$$PU = 0.159*F27 + 0.290*F29 + 0.176*F31 + 0.431*EOU + \varepsilon. \quad (10)$$

Dependent Variable ATT. Although the VIF values obtained in an early regression permitted to validate the orthogonality of the independent variables, the variable p_value EOU was very high, meaning the little importance of this variable in the formation of the variable ATT. Then it was carried out a new linear regression, removing this variable from the model. The adjusted model, including only the variable PU, explains 26.6% of the variability of ATT. Using the unique Standardized Beta Coefficient results the following structural equation for ATT:

$$ATT = 0.517*PU + \varepsilon. \quad (11)$$

Dependent Variable USE. Since now there is only one independent variable, it was carried out a simple linear regression. The fitted model explains 30.8% of the variability of the variable USE and the structural equation for this variable is as follows:

$$USE = 0.556*ATT + \varepsilon. \quad (12)$$

Confidence Intereval. With the structural equations were calculated the values of EOU, PU, ATT and USE and also the mean, the variance and the maximum of the variable USE. These values are found in the table 3.

Table 3. KMO and Bartlett's Test values

Mean	1.6926
Variance	0.066
Maximum	2.05

The values of the mean and variance enable to determine the 95% confidence interval for the mean of the variable USE. Computing the lower and the upper limits, results the following 95% confidence interval for the mean:]1.6712, 1.7140[.

The ratios between each one of these limits and the maximum of the variable USE are 0.815 and 0.836, which shows that the probability of adoption of the researched technology is between 81.5% e 83.6%.

6 Conclusion

It should be noted that the objective of this research didn't include a new validation of the TAM. It has already been the focus of so many validations, so it was concluded that was not necessary to carry out a new one. The intention was only to use it as a theoretical referential to guide the research in order to evaluate the adoption level of a technology that can provide an effective safety of the baby.

When was made the database, it seemed that the age and the qualifications of respondents had an influence over the given answers. This fact was not investigated, since the TAM does not include these types of variables. However, their influence over an innovation adoption may be an interesting matter for a future research.

References

1. Garcia, M.M.M., Camargo, N.L.: Estudo do Desenvolvimento da Linguagem de Crianças de 0 a 2 Anos. FAHU/FAEF, SP (2007)
2. Bricks, L.F.: Choro Excessivo e Cólica em Lactentes. HC-FMUSP, SP (2001)
3. Reed, U.C.: Doenças Neuromusculares. J. Pediatr (Rio J.), porto Alegre 78(suppl.1) (2002)
4. Branco, A., Fekete, S.M.W., Rugolo, L.M.S.S.: O Choro como Forma de Comunicação de Dor do Recém-Nascido: uma Revisão. Artigo de revisão, UNESP (2006)
5. Jesus, M.B.P., Nogueira, V.O.: Diagnósticos de Enfermagem de Pacientes Pediátricos com Epilepsia: um Estudo Prospectivo. ConScientiae Saúde, SP (2008)
6. Oliveira, A.B., Oliveira, A.O., Miguel, M.D., Zanin, S.M.W., Kerber, V.A.: Hipotireoidismo sob a Ótica Farmacêutica Generalista. Visão Acadêmica, Curitiba (2002)
7. Reddy, C.M., Willoughby, L.F., Hara, S., Crump, E.P.: Neonatal Meningitis Due to Enterobacter Cloacae. Journal of the National Medical Association 70(5) (1978)
8. Naeye, R.L., Ladis, B., Drage, J.S.: Sudden Infant Death Syndrome. Am. J. Dis. Child (1976)
9. Rode, J.A., Kaye, J.J.: Is to Nurture in Technology's Nature?(s/d)
10. Reami, T.F., Oliveira, L.H.M., Ramires, E.S.: Baby Alarm. Projecto de Materiais e Produtos. Senai (2009)
11. Márquez, J.M.A., Sanguino, T.J.M., Aguilar Nieto, F.J., Barrera, J.J.C., Mateos, M.F.M.: An Image Acquiring, Processing and Transfer System over Bluetooth for an Educational Robotic Platform. In: 7th Conference,on mobile robots and competitions, Albufeira (2007)
12. Santos, S. C. B.: Reconhecimento de Voz Contínua para o Português Utilizando Modelos de Markov Escondidos. Tese de doutoramento, PUC, Rio de Janeiro (1997)
13. Pires, P.J., e Costa Filho, B.A.: Fatores do Índice de Prontido à Tecnologia (TRI) como Elementos Diferenciadores entre Usuários e não Usuários de Internet Banking e como Antecedentes do Modelo de Aceitação de Tecnologia (TAM). RAC, Curitiba (2008)
14. Mathieson, K.: Predicting user intentions: Comparing the technology acceptance model with the Theory of Planned Behavior. Information Systems Research, Michigan (1991)
15. Davis Jr., F.D.: A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. Doctoral dissertation. Sloan School of Management, Massachusetts Institute of Technology (1986)
16. Silva, P.M.: Modelo de Aceitação de Tecnologia (TAM) Aplicado ao Sistema de Informação da Biblioteca Virtual em Saúde (BVS) nas Escolas de Medicina da Região Metropolitana do Recife. Dissertação de mestrado, BC/FCMPB, João Pessoa (2008)
17. Maroco, J.: Análise Estatística com Utilização do SPSS. Edições Sílabo, Lisboa (2010)

Design of Discrete Variable Structure Controller Based on Variable Boundary Layer

Shibin Su, Heng Wang, Hua Zhang, and Wei Xiong

Robotics Laboratory, School of Information Engineering
Southwest University of Science and Technology
MianYang, SiChuan, China
{gongjuqianzi,wh839}@163.com

Abstract. Variable structure control is recognized as the most efficient method for control of uncertain systems. In order to reduce the chattering and improve the stability of the Discrete-time Variable Structure Control System (DVSCS) with parameter uncertainties and external disturbances, this paper proposes a variable structure controller based on variable boundary layer. Simulations results demonstrate the controller is not only effectively reduce chattering of the DVSCS, but improve the asymptotic stability and robustness of system. In addition, the results also revealed that the controller has strong effectiveness and robustness of the trajectory tracking.

Keywords: DVSCS, chattering, variable boundary layer.

1 Introduction

Variable structure controller with sliding mode control was first proposed in the early 1950's. Through Most mathematical models of the actual system are not precision because the actual systems contain nonlinearity and uncertainty inevitably, variable structure control is recognized as the most efficient method for control of uncertain systems. And due to these advantages and simplicity of implementation, variable structure controllers have widely been used in various applications. However, switching can produce chattering in practice. The chattering phenomenon will excite unmodeled high frequency dynamics, which may form a serious disturbance and cause the system unstable. The chattering phenomenon is one of the current problems in modern sliding mode control theory, especially in the DVSCS with parameter uncertainties and external disturbances. Therefore, in recent years, Reducing chattering of variable structure control attracts control domain's attentions. Many approaches for reducing the chattering have been proposed.

Wu Han-song [1] presented a novel variable reaching law based on the system state bound norm. Then stability condition of Variable speed reaching law was given in theory. In [2], the quantitative mathematical relationship between the steady-state error of the DVSCS and the saturation characteristics of the variable thickness of the boundary layer was presented. In [3], Su Ruixiang allowed the movement of system state tends to the origin step by step through the attenuation coefficient to smaller boundary layer. The stability of system was improved. Reference [4] presented a new

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 71–78, 2010.

chattering free support vector regression sliding mode control law based on Linear Matrix Inequalities. Reference [5] presented sliding mode control synthesis based on the observation, which robust against sufficiently small delay variations and external disturbances. Li Yan [6] raised a variable compression coefficient reaching law, which making the system state points obtain a different convergence speeds according to their distance from the near and far away from the sliding surface. This method made the chattering and the amplitude of the system switching control significantly reduced. Zhu Qidan [7] reduced the chattering and improved effectively robustness of the system through the improvement reaching law of discrete variable structure control. Mu Lijun [8] presented a tracking method based on sliding mode control strategy, which making the reaching problem transform into the tracking problems. She achieved the interference suppression by the state feedback control law of error system and the interference model based on internal model principle. R.Benayache [9] improved the tracking performance of a nonlinear system using varying boundary layers instead of fixed boundary layers; others see [10–14].

Many traditional design methods reduced chattering of variable structure control. But there are few methods consider the DVSCS with parameter uncertainties and external disturbances. Such as, Lingfei Xiao proposed the multi-step sliding mode prediction model, which included the function of compensating for system parameter perturbation and external disturbance. Under the control law, closed–loop systems chattering had been eliminated effectively [15]. In this paper, a solution is proposed based on variable boundary layer in order to solve the chattering problem in the DVSCS with parameter uncertainties and external disturbances. The thickness of the boundary layer is changed with the system state position changing. Theory and simulation analysis show that, this method not only can make the system state trajectory smooth sliding surface and eventually enter into the sliding surface.

2 Variable Boundary Layer Analysis

2.1 Model Analysis

Consider the form of the DVSCS with parameter uncertainties and external disturbances as follow

$$x(k+1) = (A + \Delta A)x(k) + bu(k) + f(k) \tag{1}$$

Where $b > 0$, and ΔA, f represent system parameter uncertainties and external disturbances. Here, assume ΔA and f meet the match condition in the DVSCS with parameter uncertainties and external disturbances $\Delta A = b\tilde{A}, f = b\tilde{f}$.

Thus Eq. (1) reduces to

$$x(k+1) = Ax(k) + b[u(k) + h(k)] \tag{2}$$

where $h(k) = \tilde{A}x(k) + \tilde{f}(k)$ consists of system parameter uncertainties and external disturbances. The sliding mode switching function is defined as

$$s(x) = c^T x(k) \tag{3}$$

Where $c^T = [c_1 \cdots c_n]$. The choosing of $c_i (i = 1, \cdots, n)$ should guarantee the stability and dynamic performance of ideal quasi-sliding mode and $c^T \neq 0$. The exponential reaching law for discrete-time systems is constructed as follow

$$s(k+1) = s(k) + T(-\varepsilon \operatorname{sgn}(s(k)) - qs(k)) = (1 - qT)s(k) - \varepsilon T \operatorname{sgn}(s(k)) \quad (4)$$

Then according Eq. (2) and Eq. (3), $s(k+1)$ can be transformed into form as follow

$$s(k+1) = c^T [Ax(k-1) + b[u(k-1) + h(k-1)] \quad (5)$$

For Eq. (4), with an initial value $s(0) \neq 0$, if $k \to \infty$, then $| s(k) | \to \dfrac{\varepsilon T}{2 - qT}$;

if $| s(k) | = \dfrac{\varepsilon T}{2 - qT}$, then $s(k) = -s(k-1)$. Then it can be learn that, if $|(s(k)| \geq \phi$, $s(k+1)$ will gradually reduce, until entering into the boundary layer. If $|(s(k)| < \phi$, $s(k+1)$ will remain in the boundary layer [11].

The above analysis shows that the exponential reaching law guides the system state into the boundary layer. However, the system state cannot eventually be guided to sliding surface. State trajectory movement in the boundary layer is often non-stop campaign across the sliding surface, doing high-frequency chattering, which is harmful to the system [12].

2.2 Variable Boundary Layer

The boundary layer thickness is significant for the variable structure controller. Fixed boundary layer thickness is no longer to make the system state to zero, but a small neighborhood around zero. In the small neighborhood, the system state point does high-frequency chattering.

In order to not only solve the problem and gain a better compromise between chattering and better tracking precision in the presence of the DVSCS with parameter uncertainties and external disturbances, here, the variable boundary layer function is presented, which can be defined as follow

$$\phi(k) = \begin{cases} \phi_0 & ,| s(k) | > \phi_0 \\ \phi_0 e^{-kt} & ,| s(k) | \leq \phi_0 \end{cases} \quad (6)$$

where ϕ_0 is a small constant , but greater than zero.

The variable boundary layer function $\phi(k)$ s shows that, if $| s(k) | > \phi_0$, the state trajectory in the boundary layer outside, $\phi(k) = \phi_0$, the boundary layer thickness is a constant value ϕ_0 . Then reaching law ensure state enter into sliding mode area with a smooth trajectory. If $| s(k) | \leq \phi_0$, the boundary layer thickness is defined by $\phi(k) = \phi_0 e^{-kt}$. As the result of negative exponential function as the coefficient, the boundary layer thickness $\phi(k)$ will be lead to zero. Thus, the amplitude of $| s(k) |$

will quickly decay，and eventually converge to zero. The system state trajectory be directed to the sliding surface, and remain in the sliding surface.

From the paper [2], we can get that the boundary layer thickness $\phi(k)$ s should be bound by the following equation

$$\dot{\phi}(k) = -r\phi(k) + k(x_d, k) \tag{7}$$

where $r = \dfrac{\hat{k}(xd,k)}{\phi(k)}$, and $r > \dfrac{\max\limits_{\forall k}(F(xd,k) + D(k))}{\lambda^{n-1} e_i}$, e_i, is the value of system steady-state error.

In the actual industrial control system, control and observation is a pair of dual relationship. For variable structure control for discrete-time system with parameter uncertainties and external disturbances, disturbance observer are gotten from the reference [13] as follows

$$u(k) = -\hat{h}(k) + (c^T b)^{-1}[cTx_r(k+1) - c^T Ax(k) + qs(k) - \eta sat(\frac{s(k)}{\phi(k)})] \tag{8}$$

$$\hat{h}(k) = \hat{h}(k-1) + (c^T b)^{-1} g[s(k) - qs(k-1) + \eta sat(\frac{s(k-1)}{\phi(k)})] \tag{9}$$

where $\tilde{h}(k) = h(k) - \hat{h}(k)$, η, q, g are small constants and greater than zero. In order to illustrate smooth switching across the surface$s = 0$, the switching control law in Eq. (8) is replaced by a sat function instead of the sign function.

Simultaneously, two theorems are gotten as follows:

Theorem 1 There are two lemmas gained by the sliding-mode controller and disturbance observer can be formulated as

$$\begin{aligned} s(k+1) &= qs(k) - \eta sat(\frac{s(k)}{\phi(k)}) + c^T b\tilde{h}(k) \\ \tilde{h}(k+1) &= (1-g)\tilde{d}(k) + h(k+1) - h(k) \end{aligned} \tag{10}$$

Theorem 2 $\forall m > 0$, if$|h(k+1) - h(k)| < m$,

$\forall k_0$, while, $k > k_0$, $\tilde{h}(k) < m / g$, where $0 < g < 1$.

The two theorems can be easily proved.

3 Robust Stability Analysis

For the variable structure controller Eq. (5), if the following 1)--3) hold, then the system is stable:

1) $0 < q < 1, 0 < g < 1, q\phi(k) > \eta$;

2) $|h(k+1) - h(k)| < m$ (holds for all k for some constant; $m > 0$;

3) $\eta > c^T b \dfrac{m}{g}$, and. $\eta > |c^T b\tilde{h}(k)|$.

Proof:

According to theorem 2, if$|s(k)| < \phi(k)$, then $s(k+1)$ remain in the boundary layer. If $|s(k)| \geq \phi(k)$, divide two kinds of case discussion reach conditions:

(1) If $s(k) \geq \phi(k)$,

$$s(k+1) = qs(k) - \eta + c^T b\tilde{d}(k) < qs(k) < s(k)$$

$$s(k+1) - s(k) < 0$$

as $q\phi(k) > \eta, \phi(k) - \eta > q\phi(k) - \eta > 0$,

$$s(k+1) + s(k) = (q+1)s(k) - \eta + c^T b\tilde{h}(k) \geq (q+1)\phi(k) - \eta + c^T b\tilde{h}(k)$$

$$= q\phi(k) + \phi(k) - \eta + c^T b\tilde{h}(k) > \eta + c^T b\tilde{h}(k) > 0$$

thus $s(k+1)^2 < s(k)^2$.

(2) If $s(k) \leq \phi(k)$,

$$s(k+1) = qs(k) + \eta + c^T b\tilde{h}(k) > qs(k) > s(k)$$

$$s(k+1) - s(k) > 0$$

$$s(k+1) + s(k) = (q+1)s(k) + \eta + c^T b\tilde{h}(k) \leq -(q+1)\phi(k) + \eta + c^T b\tilde{h}(k)$$

$$= -q\phi(k) - \phi(k) + \eta + c^T b\tilde{h}(k) < -\eta + c^T b\tilde{h}(k)$$

$$< 0$$

thus $s(k+1)^2 < s(k)^2$.

The robust stability analysis states that each switching function remains smaller than regardless of the size of the disturbance, if the conditions 1)--3) hold. When the disturbance varies slowly, is small and accordingly small is sufficient to satisfy the condition 3). These make small, thus forcing the switching function to remain near zero. The chatter is inevitable because of the discontinuous control. However, $\phi(k)$ ultimately tends to zero, making the system state point tends to the origin. The system state trajectory is directed to the sliding surface, thereby enhancing the system's stability and robustness are gotten enhanced.

According above analysis, the improvement variable structure controller still meets the conditions of entry conditions sliding mode area

$$|s(k+1)| < |s(k)|.$$

And, the reaching law continues to meet the reaching conditions of the system state [14]

$$\lim_{s_k \to 0} s_k \dot{s}_k \leq 0, k = 1, \ldots, m.$$

The stability analysis shows that the method proposed based on variable boundary layer is still able to meet the system stability. Firstly, the system state point enters into the preset boundary layer. With the boundary layer decreasing, it eventually enters into the sliding mode surface. The diagram is as follow.

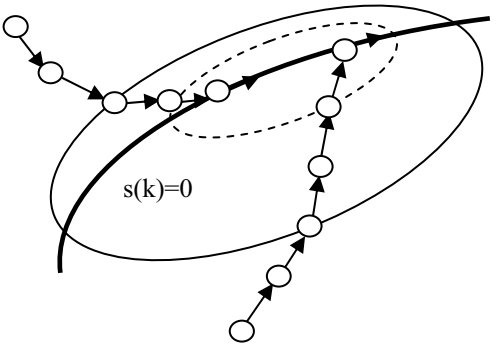


Fig. 1. Gait diagram of switching function

4 Simulation

Consider the variable structure control mode for the nonlinear DVSCS with parameter uncertainties and external disturbances as follow

$$x(k+1) = Ax(k) + b[u(k) + h(k)]$$

where initial condition $A = \begin{bmatrix} 1.3 & 0.001 \\ 0 & 0.783 \end{bmatrix}$, $b = \begin{bmatrix} 0.0011 \\ 0.153 \end{bmatrix}$, assume system instructions sinusoidal signal $x_r(k) = 0.5\sin(4\pi t) + 0.3\sin(7\pi t)$, sliding mode parameter $c^T = [15\ 1]$, system initial state $x_0 = [1.5\ \ 0]$, $q = 0.8$, $\phi = 0.05$, $g = 0.94$, $m = 0.01$, disturbance signal $h(k) = \tilde{A}x(k) + \tilde{f}(k) = 0.7\cos(4\pi t)$.

Computer simulation results are as follows. The Fig. 2 is simulations results of system with constant boundary layer, and the Fig. 3 is simulations results of system using variable boundary layer.

Comparing the Fig. 2 and Fig. 3, it can be easily seen that, the tracking effect of the system input and output using variable boundary layer method is improved effectively. It's in 0.3s that system state point enters into the stable tracking. Simultaneously, the system's tracking accuracy is improved. It has a good practical significance. Obviously, the system is asymptotical stable, and sliding mode motion trends to the origin point in finite time. The simulation analysis shows that the rapidness and robustness of the system are improved. Simulation results show that the controller eliminates the chattering phenomenon and obtains good tracking performance.

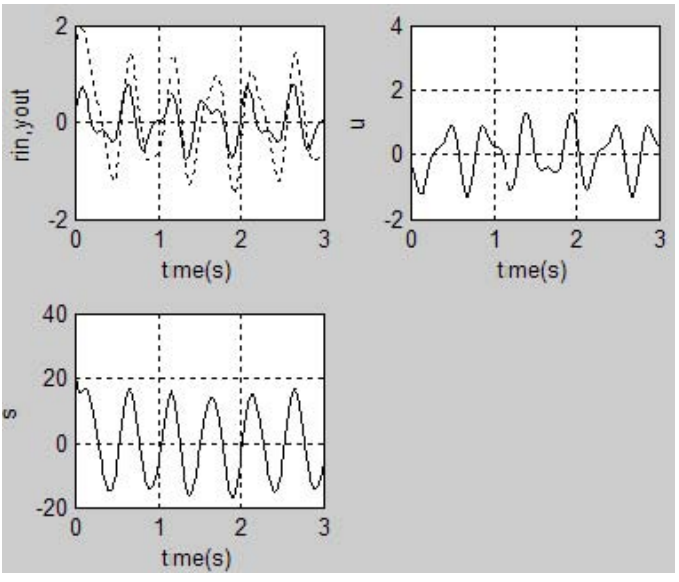


Fig. 2. Using constant boundary layer

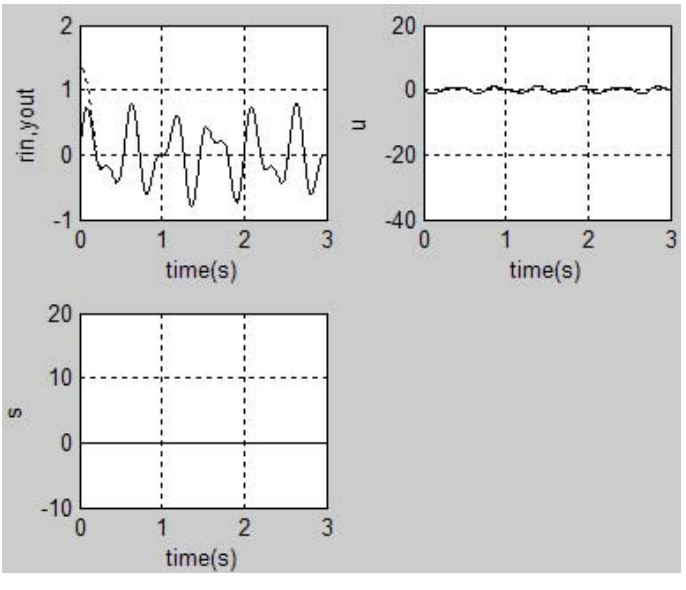


Fig. 3. Using variable boundary layer

5 Conclusions

In this paper, an approach based on variable boundary layer is presented. It can be easily conclude that, the sliding mode control with variable boundary layer is much

better than the conventional sliding control with constant boundary layer. The obtained results prove the viability of this control method and presented good performances in term of robustness to the DVSCS with parameter uncertainties and external disturbances. The variable structure controller laid a foundation for the further study of discrete-time systems with parameter uncertainties and external disturbances.

References

1. Hansong, W., Lizhong, S., Qiong-hui, Y.: Chattering analysis and robustness study of discrete variable rate reaching law. Journal of Naval University of Engineering 16, 41–44 (2004)
2. Zhao-qiang, L., Deyun, Z.: Design of chattering-free global sliding-mode discrete Variable Structure Control algorithm. Computer Engineering and Applications 44, 226–227 (2008)
3. Ruixiang, S., Lizhong, S.: Discrete Variable Structure Control Strategy With Decreasing Quasi-Sliding Mode Band. Computing Technology and Automation 20, 54–58 (2001)
4. Li, L., Li, J.: Chattering-Free Support Vector Regression Sliding Mode Control. In: 2008 International Conference on Computational Intelligence and Security, vol. 99, pp. 231–234 (2008)
5. Orlov, Y., Perruquetti, W.: Sliding Mode Control Synthesis of Uncertain Time-Delay Systems. Asian Journal of Control 5, 568–577 (2003)
6. Yan, L.: Variable Structure Control for Discrete-time Systems and Its Simulations (2008)
7. Qidan, Z., Tong, W.: New variable structure control scheme for discrete-time systems. Control and Decision 24, 1209–1213 (2009)
8. Lijun, M., Cunchen, G.: Sliding mode tracking control for a kind of discrete-time systems with time-delay and disturbance. Control and Decision 8, 874–878 (2008)
9. Benayache, R.: Design and implementation of sliding mode Controller with varying boundary layer for a coupled tanks system. In: 17th Mediterranean Conference on Control & Automation, vol. 7, pp. 1215–1220 (2009)
10. Boiko, I., Fridman, L., Castellanos, I.M.: Analysis of second-order sliding-mode algorithms in the frequency domain. IEEE Trans. Autom. Control 49, 946–950 (2004)
11. Jinkun, L.: MATLAB Simulation for Sliding Mode Control. Tsinghua University Press, Beijing (2005)
12. Lee, H.: Chattering Suppression in Sliding Mode Control System. The Ohio State University (2007)
13. Eun, Y., Kim, J.-H.: Discrete-Time Variable Structure Controller with a Decoupled Disturbance Compensator and Its Application to a CNC Servomechanism. IEEE Transactions on Control Systems Technology 7, 414–423 (1999)
14. Li, Y.-F.: High Precision Motion Control Based on a Discrete-time Sliding Mode Approach (2001)
15. Xiao, L., Su, H.: Multi-step Prediction Based Discrete-time Sliding Mode Control Algorithm. In: Second International IEEE Symposium on Intelligent Information Technology Application, pp. 731–735 (2008)

Cognitive Biometrics: Challenges for the Future

Kenneth Revett[1] and Sergio Tenreiro de Magalhães[2]

[1] University of Westminster
School of Electronics & Computer Science London, England
revettk@westminster.ac.uk
[2] Universidade Catolica Portuguesa
R. de Camoes
4710-362 Braga
stmagalhaes@braga.ucp.pt

Abstract. Cognitive biometrics is a novel approach to user authentication/identification which utilises a biosignal based approach. Specifically, current implementations rely on the use of the electroencephalogram (EEG), electrocardiogram (ECG), and the electrodermal response (EDR) as inputs into a traditional authentication scheme. The scientific basis for the deployment of biosignals resides principally on their uniqueness -for instance the theta power band in adults presents a phenotypic/genetic correlation of approximately 75%. The numbers are roughly the same for ECG, with an heritability correlation for the peak-to-peak (R-R interval) times of over 77%. For EDR, the results indicate that there is approximately a 50% heritability score (h^2). The challenge with respect to cognitive biometrics based on biosignals is to enhance the information content of the acquired data.

Keywords: Cognitive biometrics, EEG, ECG, EDR, user authentication.

1 Introduction

Cognitive biometrics utilises a biosignal approach to user authentication. These systems utilise biological based signals such as the electrocardiogram (ECG), the electroencephalogram (EEG), and the electrodermal response (EDR) as the inputs to an authentication system. These signals are generated by the heart, brain, and theautonomic nervous system respectively – and are recorded using standard equipment in a generally non-invasive fashion. The basic approach is to record these signals – and use them directly – as a biometric signature. Each of these biosignals presents a wealth of information that can be extracted quite easily using a single recording system, such at the NeXus-4 system (www.nexus.com), which provides 4 channels for recording a combination of EEG, ECG, or EDR, using wireless technology for data transport to a server for authentication purposes. How these biosignals can be harnessed for user authentication (andidentification) is addressed in this paper, starting with the historical development of the ECG for user identification.

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 79–86, 2010.

2 Background

The use of the ECG as a biometric was examined in detail by Forsen in 1977 , a prescient paper that also discussed the deployment of EEG as a biometric tool [1]. The ECG records the electrical activity generated by the beating heart – generating a characteristic waveform which is depicted in figure 1. This technology has a long and venerable history, beginning officially in 1887 [2]. The heart utilizes electrical activity to activate the muscles required to pump blood through the circulatory system. By placing sensitive recording electrodes at particular regions around the heart – the signals can be detected. The signals generated by the heart beat forms a regular pattern (see figure 1) that records the electrical activity of the heart. This signal was utilized by Forsten in an attempt to determine the individuality of the ECG – if it was determined that the signal is unique – he proposed that this would serve as a useful biometric technique.

In Forsen's approach, the recording of the ECG was accomplished in a very non-invasive fashion – he utilized two electrodes that were attached to the index fingers without the use of a ground lead or electrode paste. Data was collected from subjects at three sessions of 30-40 seconds each. The data was filtered with a 3 KHz cutoff frequency and the data was digitized for subsequent analysis (for more details consult [1]. Several features were extracted for subsequent classification purposes. A total of 10 features were utilized: five time intervals and five amplitude differences. The time points correspond to the 5 major deflection points in the signal (labeled P,Q,R,S, & T). The amplitude measurements we produced using the same five time point fiduciaries, with the addition of a 6^{th} halfway between S and T deflection points. These features are utilized to produce a reference vector for the individual. When the same user requests authentication, several heart beats are recorded (takes a few seconds only), and the average of the authentication request trials is comapresd with the reference vector. The results of this approach, based on type I and type II errors were extremely encouraging, yielding values of 1.2% and 1.1% respectively. This is a phenomenal result – considering the small number of features utilized.

The results from the Forsen study have been confirmed by other researchers. Silva and colleagues published results indicating a successful classification rate of over 99% from a collection of 26 subjects, using a contingency matrix analysis approach [3]. Also note that in the Silva study, only just over 1 minutes worth of ECG recording was utilized for this high classification accuracy (63 seconds). A study by Israel and colleagues examined the stability of the ECG as a biometric modality [4]. Their results indicate that the features extracted for classification purposes were independent of sensor location, invariant to the individual's state of anxiety, and unique to an individual. There are several other studies that employ ECG as a method of user identification, most of which provide exceptional classification results (greater than 95% accuracy – and many reaching 100% (see [5][8]). This is a desirable quality, as the stability of the signal must be sufficient for robust classification. This issue of stability forms a central research question in the next topic: the electroencephalogram (EEG).

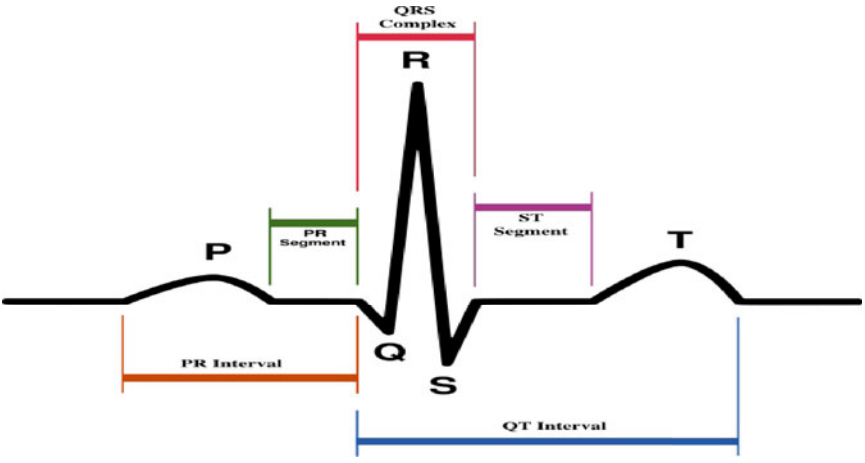


Fig. 1. A typical ECG pattern for a single heart beat, with indications of the principle fiduciary marks

As previously mentioned, Forsen proposed that EEG signals could be deployed for user verification [1], though Vidal may be the first to report on the possibility of a brain-computer interaction [9]. Indeed, there are a number of studies that have investigated the use of EEG data in this capacity [10]-[14]. In a wider sense, EEG signals have been used as a means of allowing handicapped individuals interact with computer systems, in what is termed the brain- computer interface. Both of these issues are directly related and form a very exciting area of research, much of which is applicable to behavioral biometrics. First a brief discussion of the basis of the EEG is in order.

The EEG is a signal that is generated by the collective activity of neuronal generators. That is, brain activity produces an electrical signal that can be recorded by placing sensitive electrodes on the surface of the scalp (see Figure 2). What is required for the signal to be recorded at the scalp is a collection of neurons firing synchronously, and oriented towards the surface of the head. Provided these conditions are met, a stereotyped signal is recorded from an array of electrodes positioned over the entire surface of the scalp. As suggested by figure 8.4, a tremendous amount of data generated during and EEG recording. Typically, anywhere from 18-256 electrodes are positioned on the scalp, each providing a time series sampled at 0.5-1.0 KHz. Typically, this generates hundreds of megabytes of data that must be analyzed in order to extract useful information. Therefore, the signals that are generated from EEG have to be extensively preprocessed before they can be utilised – and tools such as EEGLAB are remarkably well suited to perform this task. The extent of the pre-processing is contingent upon the task at hand though. EEG as a cognitive psychological tool typically requires the isolation of various components within the EEG signal (see figure 2 for examples of EEG signals). That is, the brain is continuously and spontaneously active. The firing of neurons, which is an electrical process, generates recordable signals that form a background, upon which is superimposed the activities of specific collections of neurons that respond according to the engagement of a variety of cognitive tasks, such as reading, thinking of

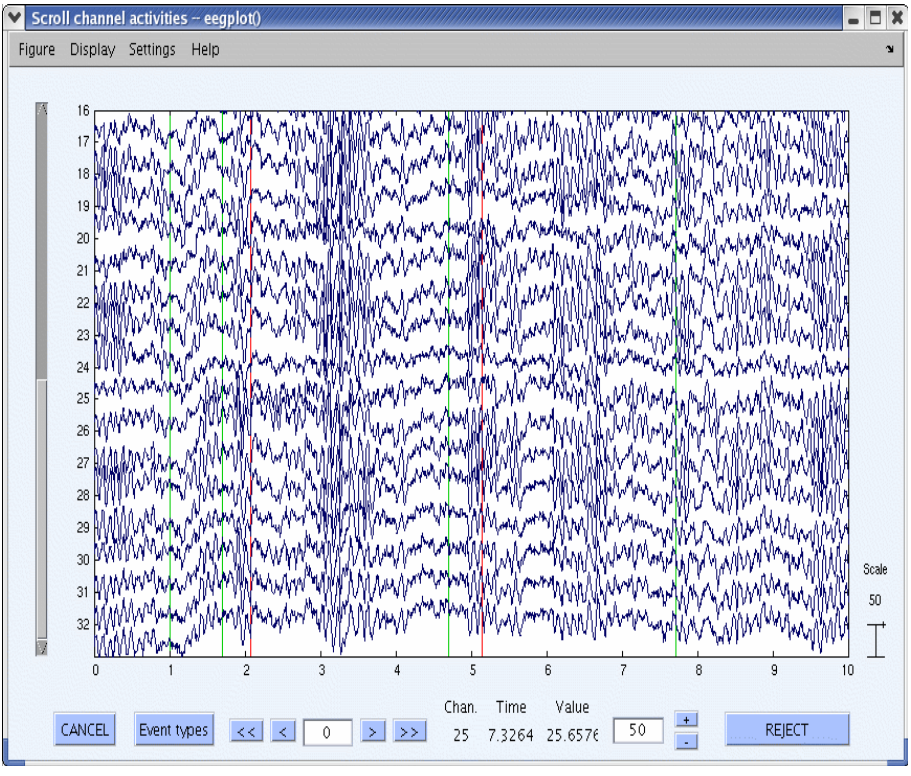


Fig. 2. Screenshot from the EEGLAB suite of utilities for processing EEG data. The x-axis is time (in seconds) and the y-axis is the raw signal amplitude (in v),μ measured at a collection of electrodes. Taken from EEGLAB [15].

an image, vocalization etc. The difficulty is identifying specific responses within the ongoing background activity, indicated as a collection of EEG signals, which are reliably correlated with the engagement of a specific cognitive task. Part of the difficulty in performing this task is the issue of spatial resolution. One question to be answered using this technique is where within the brain did a particular identified signal emanates from? The goal here is to associate an anatomical region of the brain with a response – thereby associating function with anatomy. A huge array of researchers has been engaged in this effort, and the research output has been phenomenal. A consistent finding from this research effort is that there are a number of fundamental changes that occur (are superimposed upon) as a result of cognitive processing.

One of the key findings from EEG research is the discovery of an event termed the "P3." The P3 (or P300 as it is also termed) is an example of an event related potential (ERP) [16]. An ERP is a signal that indicates a cognitive event has occurred – the P3 is typically produced in what is termed an "odd-ball" paradigm. This is a scenario in which a subject is presented with a stimulus to identify and a stimulus not related to the expected stimulus is presented. In addition, there are a variety of other characteristic signals that are produced by the brain in response to typical stimuli [17],[18].

These stimulus evoked responses have been employed as a means of facilitating the interaction between humans (and animals as well) and computers – the essence of the brain-computer interface.

In an interesting paper entitled "Pass-thoughts, authenticating with our minds," the notion that we may be able to authenticate – for instance entering a password simply by thinking of the password [19]. The basis of this work is that through the deployment of a BCI based mechanism, which relies on the use of EEG acquired signals from a subject. The research question of interest here is the individuality and reliability of these signals (i.e. ERPs and/or related events). These questions can be addressed by examining the data generated by appropriately controlled experiments.

Paranjape and colleagues have successfully used this technology to identify a set of 40 subjects [20]. In their work, the subject identification scheme was 100% accurate for training cases, and 80% accurate for test case. Polous and colleagues were able to accurately identify subjects (classification accuracy between 80-100%) [21]. Riera and colleagues provide data that yields a true acceptance rate (TAR) of 95+%, with an EER of 5.5% [11]. These results indicate that this technology, as a means of user identification *per se*, is reasonably accurate – though not yet ready for commercial distribution. This research area is still at an early stage developmentally, and it is expected that this technology will improve in the near future, yielding very promising results.

The electrodermal response (EDR) can be measured quite easily using the same technology deployed in EEG and ECG, providing a unified signal acquisition system. A typical EDR signal is presented in Figure 3, which displays a 60-second recording from a single subject. Essentially, EDR simply measures the electrical resistance between two points. The resistance of the skin will change due to the emotional state of the subject, which is controlled by the autonomic (sympathetic) nervous system. The autonomic system in turn controls the activity of sweat glands which are embedded in the middle layer of the skin. The stability and heritability of the EDR has been studied in a large cohort of monzygotic and dizygotic twins([22].

The results of this study indicate that approximately 50% of the variance between individuals can be accounted for by a single latent phenotype. Although there is little published data on the deployment of EDR specifically as a biometric directly, there are reports that indicate that EDR can be used to acquire information about the emotional state of an individual (see [23]). Knowledge of the emotional state can provide adjunctive information that could be utilised for authentication purposes. Clearly, an obvious use would be to determine if a person feels nervous during their authentication attempt (for a discussion on emotional states see [24]). This might signal an alert which could then be used to inform other biometrics - in addition to being deployed directly.

These various biosginals each provide a significant level of user authentication and even identification in many cases. When combined into a multi-modal approach, the authentication/identification levels approach those of physiological based biometrics (such as fingerprints and iris scanners). The HUMABIO (www.humanbio-eu.org) system is an example of such an approach. This is a total authentication system employing a virtual reality engine, EEG analysis, and a host of other behavioral and physiological based biometric technologies. Figure 4 presents a component that is the heart of their

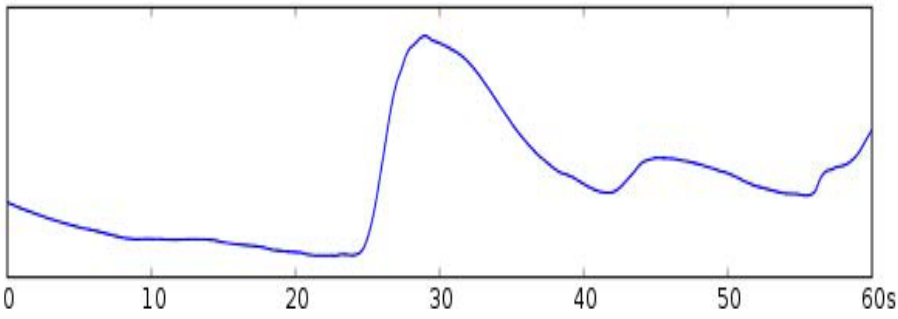


Fig. 3. A sample of the EDR measured over a 60 second period using standard Ag-AgCl electrodes placed on the palmar surface. Taken from (http://en.wikipedia.org/wiki/Galvanic_skin_response)

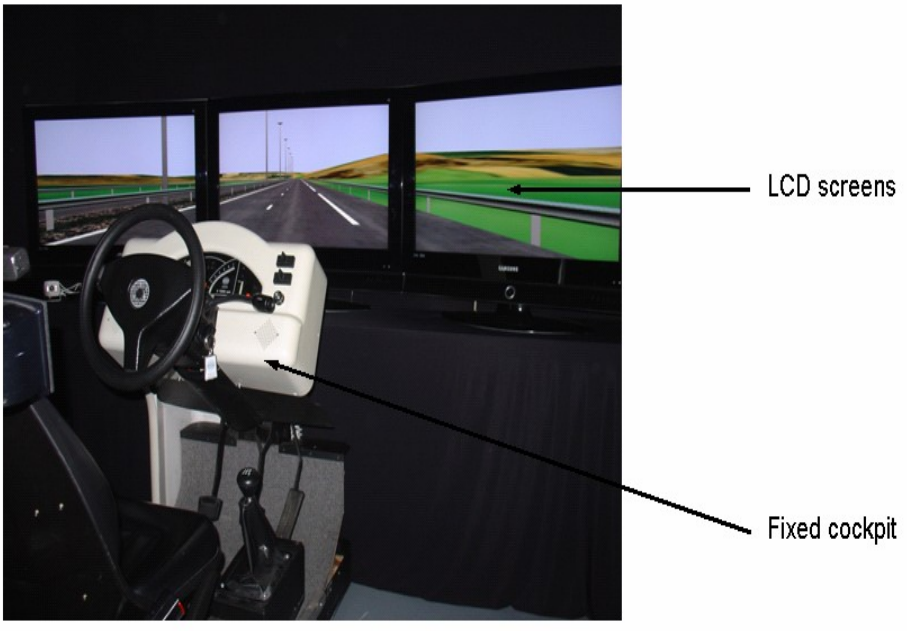


Fig. 4. The traffic simulator – the heart of the HUMABIO authentication system (www.humabio-eu.org). This system employs both EEG and ECG that records users responses and builds a user profile which may be used for subsequent user authentication.

simulator, which they term "traffic model." This technology is a state-of-the-art approach to user verification – which utilizes a variety of approaches such as indicated below in their traffic simulator. This system acquires EEG and ECG data and fuses them for subsequent user verification. It is hoped that this is one of many such systems that will be developed in the near future.

3 Conclusion

This paper highlights the deployment of biosignals as the basis for a novel approach to biometrics, termed 'cognitive biometrics.' Cognitive biometrics deploys biosignals such as EEG, ECG,and EDR as the input to an authentication system which can be used individually or in a multi-modal biometric system. These signals can be acquired using a single technology, electrodes placed on the appropriate place on the body surface (head and arms). Typically, the required amount of data can be acquired typically within 1 minute -and is painless. user acceptance issues may need to be addressed - but scalp based EEG electrodes, which typically require the use of conductive gels are being replaced through dry-electrodes, which do not require conductive gels, and are therefore easier to apply. Cognitive biometrics can record both the cognitive and emotional state of an individual -which are difficult to forge, and certainly can not be lost. They can be deployed for both static and continuous based biometrics, and can certainly be applied in conjunction with behavioral biometrics such as keystroke dynamics. These are issues that need to be addressed by the biometrics community at large.

References

1. Forsen, G., Nelson, M., Staron, R.: Personal attributes authentication techniques. In: Griffin, A.F.B. (ed.) Rome Air Development Center report RADC-TR-77-1033. RADC, New York (1977)
2. Waller, A.D.: A demonstration on man of electromotive changes accompanying the heart's beat. J. Physiol (Lond.) 8, 229–234 (1887)
3. Silva, H., Gamboa, H., Fred, A.: Applicability of lead V2 ECG Measurements in Biometrics. In: Proceedings of Med-e-Tel 2007, Luxembourg (April 2007)
4. Israel, S., Irvine, J., Cheng, A., Wiederhold, M., Wiederhold, B.: ECG to identify individuals. Pattern Recognition 38(1), 133–142 (2005)
5. Biel, L., Petterson, O., Stork, D.: ECG analysis: a new approach in human identifiaction. IEE Transactions on Instrumentation and Measurement 50(3), 808–812 (2001)
6. Kyoso, M., Uchiyama, A.: Development of an ECG identifcation system. In: Proceedings of the 23rd Annual International IEEE Conference on Engineering in Medicine and Biology Society, Instanbul, Turkey, pp. 3721–3723 (2001)
7. Kyeong-Seop, K., Tae-Ho, Y., Jeong-Whan, L., Dong-Jun, K., Heung-Seo, K.: A Robust Human Identification by Normalized Time-Domain Features of Electrocardiogram. In: IEEE EMBS 27th Annual Uinternational Conference of Engineering in Medicine and Biology (EMBS 2005), pp. 1114–1117 (2005)
8. Mehta, S.S., Lingayat, N.S.: Comparative study of QRS detection in single lead and 12 lead ECG based on entropy and combined entropy criteria using support vector machine. Journal of Theoretical and Applied Information Technology, 8–18 (2007)
9. Vidal, J.: Toward direct brain-computer communication. Annual Review Biophys. Bioeng., 157–180 (1973)
10. Palaniappan, R.: Multiple mental thought parmetric classification: a new approach for individual identification. International Journal of Signal processing 2(1), 222–225 (2005)
11. Marcel, S., del Millan, R.: Person autentication using brainwaves (EEG) and maximum a posteriori model adaptation. IEEE transcations on pattern Analysis and machine Intelligence, Special issue on Biometrics (2006)

12. Riera, A., Soria-Frisch, A., Caparrini, M., Grau, C., Ruffini, G.: Unobtrusive biometric system based on electroencephalogram analysis. EURASIP Journal on Advances in Signal Processing (2007)
13. Bell, C.J., Shenoy, P., Chalodhorn, R., Rao, R.P.N.: An image-based brain-computer interface using the P3 response, UWCSE Tech. Report # 2007-02-03, University of Washington, Seattle, Washington, USA (2007)
14. Riera, A., Soria-Frisch, A., Caparrini, M., Grau, C., Ruffini, G.: Unobtrusive biometric system based on electroencephalogram analysis. EURASIP Journal on Advances in Signal Processing (2007)
15. Delourme, A., Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience Methods 134, 9–21 (2004)
16. Sutton, S., Braren, M., Zubin, J., John, E.R.: Evoked potential correlates of stimulus uncertainty. Science 150, 1187–1188 (1965)
17. Nykopp, T.: Statistical modelling issues for the adaptive brain interface, MSc Thesis, Helsinki Univeristy of Technology (2001)
18. Gupta, C.N., Palaniappan, R.: Enhanced detection of visual-evoked potentials in brain-computer interface using gentic algorithm and cyclostationary analysis. Computational Intelligence in Neuroscience (2007)
19. Thorpe, J., Van Oorschot, P.C.: Graphical Dictionaries and the Memorable Space of Graphical Passwords. In: Proceedings of the 13th USENIX Security Symposium, pp. 135–150 (2004)
20. Paranjape, R.B., Mahovsky, J., Benedicenti, L., Kolesapos, Z.: The lectroencephalogram as a biometric. In: Canadian Conference on Electrical and Computer Engineering, pp. 1363–1366 (2001)
21. Polous, M., Rangoussi, M., Alexandris, N.: Neural network based person identification using EEG features. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal processing, vol. 2, pp. 1117–1120 (1999)
22. Crider, A., Kremen, W.S., Xian, H., Jacobson, K.C., Waterman, B., Eisen, S.A., Tsuang, M.T., Lyons, M.J.: Stability, consistency, and heritability of electrodermal response lability in middle-aged male twins. Psychophysiology 41(4), 501–509 (2004)
23. Silva, D.C., Vinhas, V., Reis, L.P., Oliviera, E.: Biometric emotion assessment and feedback in an immersive digitial environment. Int. J. Soc. Robots 1, 307–317 (2009)
24. Russell, J.A.: A circumplex model of affect. J. Personal Soc. Psychol. 39, 1161–1170 (1980)

Multimodal Biometrics and Multilayered IDM for Secure Authentication

Abdullah Rashed and Henrique Santos

R & D Centre of Algoritmi,
University of Minho,
Guimarães, Portugal
{Rashed,santos}@dsi.uminho.pt

Abstract. In the Electronic Society (e-world) users are represented by a set of data called Digital Identity (ID), which they must use for authentication purposes. Within the e-world it is certainly risky to lose the identity and this security threat must be ranking with the highest priority, forcing a solution that provides an amenable usage of digital identity. Efficient protection of the digital identity would also encourage users to enter the digital world without worries. Security needs to provide the necessary identity management (IDM) process to mitigate that threat. This paper gives an overview of IDM and suggests a framework that can be particularly useful for a secure user authentication. The proposed model appears as a multi-layered security approach, since it tries to integrate different security technologies and multimodal biometrics tools and practices, such as police, procedures, guidelines, standards and legislation. The advantages, limitations and requirements of the proposed model are discussed.

Keywords: security, digital identity management, privacy, authentication, biometrics.

1 Introduction

The growth of the internet has made it an integral part of many businesses' daily operations [7]. To enter the e-world users have to use some sort of credentials as shown in figure 1. Due to its fast and networked nature, this e-world can provide that information for non expected purposes, such as business communications and marketing [14]. Moreover, given the lack of face to face interaction, stolen or lost credentials can be easily abused to hide many types of e-crimes. To illustrate that, we will use an example provided by [19]: when users visit a bookshop they do not need to show their unique numbers or any other personal information; in contrast, when they visit e-bookshop, they have to show, at least, their IP address but normally, sites are able to capture more information. Besides, users might be fooled into providing personal digital identity to rogue sites that redirect legitimate traffic [18]. To address this issues Identity Management (IDM) is a viable solution [9] and it seems essential to protect the privacy of users in the electronic society [19] and to make them feel safe.

Digital identity is defined as the digital representation of the known information about a specific individual or organisation [3]. By definition, IDM is a set of business

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 87–95, 2010.

processes and a supporting infrastructure for creation, maintenance and use of digital identity. An IDM System (IDMS) is a system that provides the control tools for managing the identity information and the amount of it that should be available for each interaction in electronic society [19].

In order to better understand the existing risks, it is useful to have a look in the typical attacks perpetuated in the e-world against individuals.

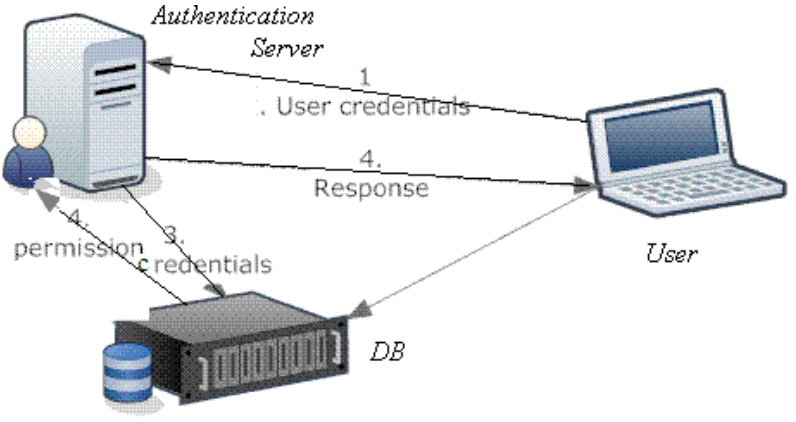


Fig. 1. Authentication process

Attacker Goals: The typical attacker tries to capture information that is confidential about a target, to gain some kind of advantage. Some examples are:

- Blackmailing: The act of extorting money by exciting to discredit or uncomfortable information would be disclosed [19].
- Revilement: publish private data to destroy victims reputation [19].
- Impersonate: stealing the identity of the victims and communicate with society with their digital identity [2].
- Denying access: when attackers obtain the identity of the victims they might change the credentials so the victims will not be able to access their information anymore [2].
- Identity attack by Phishing: the act of luring the victims to provide their digital identity to rogue websites [18].
- Attacking password: it is well known that users have many accounts (about 40) and usually use the same password; so if a rogue site could get one of them, then the secret would be broken [18].
- Privacy Attacking: disclosing private information against user willing [19].
- Attacking Databases that contain sensitive information about individuals or companies, e.g., person records, statistical databases, transaction databases, and unstructured knowledge bases [19] and [2].
- Disclose Network Anonymity: network anonymity is supposed to protect the communication traffic of a user, i.e. hide all communication [19]. When attackers gain some information they could break the anonymity. In addition, they will use this information to attack sensitive information or disclose the secrets.

IDM frameworks can help users to prevent those attacks or the awful effects they may have. But they also deal with a large set of issues concerning the privacy and other social values that are not equal for everyone in the e-world. To figure out the real extension of the IDM effect, it is useful to list the identity management dimensions [18] and [22]:

- Technical issues: concerning the infrastructure to support an IDMS.
- Legal system: especial legislation for data protection.
- Information police: for dealing with identity theft.
- Social and humanity: dealing with issues such as privacy.
- Security components: such as access control.
- Participating organizations.

We will discuss the IDM principals and overview the work done by others. This paper is organised as follows: In section 2 we overview the previous studies by a literature review. In section 3, we demonstrate our suggestion and discussion and finally, we conclude in section 4.

2 Literature Review

Ross suggested a multimodal biometrics model to improve performance, increase population coverage, deter spoofing, and facilitate indexing. They addressed a limitation of biometrics multimodal systems which is the integration. However, they thought that the integration strategies can be adopted to consolidate information [1].

Lakshmi claimed that their multimodal biometrics proposal is an efficient authentication model. Their model depended on face and fingerprint. Their approach encrypted the face images and encoded them into fingerprint images. They found that the verification accuracy was high and considered it as a cheap solution for spoofing and many other attacks. They found that their multimodal model could resist to various attacks [4].

Khan et. al. presented multimodal biometrics model for authentication. Fingerprint templates were encrypted and encoded/embedded into the face images in such a way that the features, which are used in face matching, are not significantly changed during encoding and decoding. They found that the proposed scheme is an efficient and a cheap solution [16].

The Italian Electronic Identity Card (IEID) [8], was proposed to be fully equivalent to the paper based ID card according to Italian Laws and can serve different purposes that need an authentication process.

In [11], the authors tackled the security problems of the Austrian citizen card. They showed the infrastructure, the legal aspects and many of the security techniques used. They discussed the security requirements and architecture for e-government applications. They focused on the concept of the so-called “security layer” as the core part of the security architecture.

In [19], the authors introduced some identity management techniques and their role to provide right anonymity and accountability. They overviewed the PRIME project, evaluated it and identified three fields (i.e. statistical database, network anonymity, and interactivity) from which, they believed attacks on IDM can be derived. Moreover they discussed the protection methods against these attacks.

The PRIME and PROTOTYPE projects were described in [12], where authors showed the advantages of the proposed system. They assumed that individuals can

limit the information collected about them by using pseudo-identities, certifications and cryptography when performing online transactions. Their proposed solution include negotiation between individuals and the service providers for "privacy policies" that govern how disclosed personal data can be used and which precautions must be taken to safeguard it.

In [18], online identity theft and IDM was explored and authors concluded that the Federated IDM (FIDM) model may increase the risk of identity theft. Their argument was that if users' accounts at an identity provider were successfully phished, then attackers would have opportunity to access other linked service providers.

In [17], the authors studied the intersection of international law and technology in the area of digital IDM. They discussed international treaties and guidelines concerning IDM. They discussed emerging IDM technologies and how they will enforce each other for enhancing accountability to the public.

In [23], the authors addressed the risk management in FIDM systems by presenting an identity assurance framework and its supporting technologies (as shown in figure 2). They discussed the risk mitigation framework as a part of any identity assurance solution. They demonstrated how their model based on assurance technologies could be used to report success of an identity assurance program. In addition, they discussed how their approach could be used to gain trust within a FIDM solution, both by communicating the nature of the assurance framework and mitigating risks.

A delegation model for FIDM systems was introduced in [10]. Their delegation framework is supposed to provide solutions for access control in the context of delegation. It had a function of transferring user's privileges across the entities encoded in delegation assertion extending SAML (Security Assertion Markup Language).

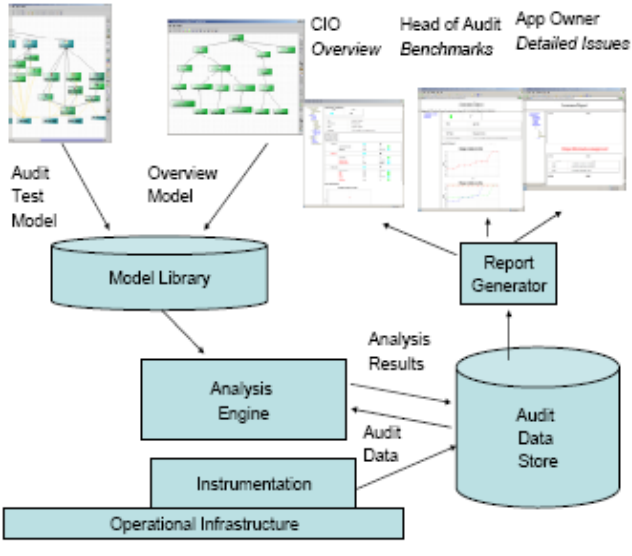


Fig. 2. IDM assurance systems [23]

An identity-based signature scheme applicable in a mediated environment was proposed in [20]. In addition, they introduced a process whereby mobile users are only required to own one private key (for each) in order to protect their communications when accessing location-based services under different pseudonyms. They assumed that an identity-based PKI would simplify significantly key management by removing the need for digital certificates. In their scheme a mediated architecture would make key revocation easier and more efficient as it allows for instant revocation of security capabilities.

In [15], the authors proposed what they called notarized FIDM model. Their model was supposed to support efficient user authentication when providers are not known to each other. "Notary service", would be introduced by a trusted third-party. They presented a cryptographic solution to prevent the leaking of secret identity information.

In [9], the authors demonstrated the importance of FIDM for internet identification. They investigated Microsoft Passport and Liberty Alliance approach concerning privacy issues in FIM, through possible business scenarios. They identified practical business scenarios and introduced their own systematic mechanisms to specify privacy preferences expression language in FIM.

In [3], the authors developed an approach to support privacy controlled sharing of identity attributes and harmonization of privacy policies in federated environments. They provided mechanisms for tracing the release of users' identity attributes within the federation. They found that the approach entailed a form of accountability since a non-compliant entity with the users original privacy preferences can be identified.

Zhou et. al. introduced spam multi-layered defense frame work. They used a combination of anti-spamming methods. They stated that their layered structure improved the filtering accuracy and reduced the number of false positives [13].

3 IDM Proposed Model

In this section, we present a multimodal biometric within multilayered IDM model (see figure 3). The assumptions of the proposed model are as follows:

- Technological tools (computer, cell-phones, hardware...) would be supplied with biometric sensors such as fingerprint, iris or odour, for user authentication purposes.
- The digital ID is biding to multimodal biometrics [6]. To overcome intrinsic biometric limitations, we suggest having multimodal biometrics:
  - Iris: a camera would be attached to the screen (mobile, laptop or desktop) so the authentication process would be done with user cooperation.
  - Fingerprint: The "enter button" would be the fingerprint scanner, to authenticate users in a stealth way, every time they press that button.
  - Odour: when buying a computer the credentials should be issued so the computer could recognize the owner, using an odour sensor. This sensor can be hidden and it would work also in stealth mode.

- Procedures:
  Users need to register at the organisation (e-resources management) and if they are eligible, they would be registered at centralized database (CDB) and get the credentials that enable them to join the organisation network. Users may do a self registration at a web site and provide their personal information and other requested data. Some privacy preferences might also be asked and stored in a personal profile. Later on, users will be allowed to change their information and preferences by user account management.
  - When users need to access some resource or to use a specific application via organisation e-resources (via user account management), they need to ask to join the organization network and do the authentication process. After that, the access control system, using information stored in the CDB, would grant their permission if they are eligible to do that.
    - For more security: we can consider all users as thin clients where the server will do everything. In this case, we achieve the security advantages:
      - Users will not need to download the SW: they just authenticate themselves whenever they need to use the resources. In addition to that, we mitigate the SW piracy.
      - It would be easy to control and analyze the traffic.

- IDM security tools:

1. Biometrics: User authentication would be done whenever they use their computers using fingerprints and face recognition.
2. Cryptography: Using cryptography was suggested by [21] and by CISCO and Intel [5].
   Public and private cryptosystems would be used for more security.
3. Policies, Guidelines, Procedures and Standards.
   -Policies
   - Users should register.
   - The SW automatically downloaded to authorized users to mitigate (SW piracy) SW distribution.
   - Digital signature using a public key cryptosystem would mitigate the impersonation.
   - Guidelines: The system would be provided with guidelines that help users understand the instructions of the system and how to use it in the correct manner: this would support the security system, and mitigate the intrusion operations.
   - Procedures: Procedures are supposed to enhance the system security as mentioned above.
   - Standards: Applying the well known standards. The system has its standard in naming and directories.
4. Legislations: There will be laws that define who is eligible and organise the penalties such as fines for those who try to break the security in this system.
5. Cryptography.
6. Access control:
   - The access will be controlled by user identity and applying access control standards.
7. Governance: CDB would be directed by the related department and the governance system should organize decision making in the system.
8. Directories: for identity and credentials

4 Conclusion

We presented IDM for a user authentication. Our proposed model was explained and assumptions are mentioned. The proposed model consists of interleaved multi-layers of security tools and techniques. Layers play an important role in defending the system effectively against the opponents attack. One advantage of the layered security system is that it can discover (detect) attacks at earlier stages, when some layers are intruded. Moreover benefits and advantages of multimodality are the ensuring of the accuracy in addition to decreasing the spoofing. Layer one concerns technologies such as firewall, router and proxy server which are placed at the demilitarized zone. Layer two concerns to legislation, governance, policies, procedures and guidelines that enhance the security and mitigate security threats. This layer is supported Layer three that represents the action that would be taken via police against those who do not respect the laws and try to misuse digital identity. Layer four implements multimodal biometrics.

Layer five concerns encryption methods, which are used for digital signature using public key infrastructure, encrypting the transmitted data via private cryptosystems. The system is covered with a layer that represents the monitor system and does different kinds of analysis to discover attackers' trials to breach the system in earlier stages. Also, this layer can help in predicting the attacker aims. It works as proactive policy: The system might block the attacker or inform police.

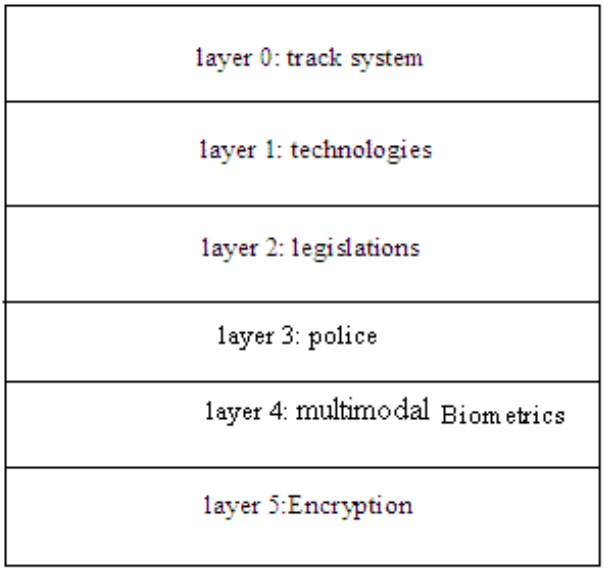


Fig. 3. layers of the proposed model

Acknowledgement

Authors would like to thank Ms. Michelle Olsen for her editing and help.

References

1. Ross, A., Jain, A.: Multimodal Biometrics: An Overview. In: Proceedings of 12th European Signal Processing Conference (EUSIPCO), Vienna, Austria, pp. 1221–1224 (2004)
2. Rashed, A.: Intelligent Encryption Decryption Systems Using Genetic Algorithms, Ph.D. Dissertation, Arab Academy, Amman, Jordan (2004)
3. Squicciarini, A., Czeskis Bhargav-Spantzel, A.: Privacy policies compliance across digital identity anagement systems. In: SPRINGL 2008, pp. 72–81 (2008)
4. Lakshmi, B., Kannammal, A.: Secured Authentication of Space Specified Token with Biometric Traits – Face and Fingerprint. IJCSNS International Journal of Computer Science and Network Security 9(7) (2009)
5. Cisco Systems, Inc., Intel: Five Myths of Wireless Networks (2006), http://www.ciscosystems.lt/en/US/.../prod_white_paper0900aecd805287fc.pdf
6. Birch, D.: Digital Identity Management. Gower (2007) ISBN: 978-0-566-08679-3
7. Taylor, D.S.: Multi-Layered Approach to Small Office Networking, the SANS Institute Reading Room site (2001), http://www.sans.org/.../hsoffice/multilayered_approach_to_small_office_networking_624
8. Arcieri, F., Ciclosi, M., Dimitr, A., Fioravanti, F., Nardelli, E., Talamo, M.: The Italian Electronic Identity Card: Overall Architecture and IT infrastructure. In: 2nd International Workshop on Certification and Security in Inter-Organizational E-Services (CSES 2004), Toulouse, France. IFIP Conference Proceedings, vol. 306, pp. 5–18 (2000)
9. Ahn, G., Lam, J.: Managing Privacy Preferences for Federated Identity Management, Workshop On Digital Identity Management. In: Proceedings of the 2005 Workshop on Digital Identity Management, Fairfax, VA, USA, SESSION: Privacy protection, pp. 28–36 (2005)
10. Gomi, H., Fujita, S.: A Delegation Framework for Federated Identity Management. In: Proceedings of the 2005 Workshop on Digital Identity Management, pp. 94–103 (2005)
11. Leitold, H., Hollosi, A., Posch, R.: Security Architecture of the Austrian Citizen Card Concept. In: 18th Annual Computer Security Applications Conference, San Diego, California, December 09 - 13, pp. 391–403 (2002)
12. Camenisch, J., Abhi, S., Sommer, D., Fischer-Hübner, S., Hansen, M., Krasemann, H., Lacoste, G., Leenes, R., Tseng, J.: Privacy and Identity Management for Everyone. In: Proceedings of the 2005 workshop on Digital Identity Management. Fairfax, VA, USA, SESSION: Privacy protection, pp. 20–27 (2005)
13. Zhou, J., Chin, W., Roman, R., Lopez, J.: An Effective Multi-layered Defense Framework Against Spam. Information Security Technical Report 12(3), 179–185 (2007)
14. Casassa Mont, M., Thyne, R.: Privacy policy Enforcement in Enterprises with Identity Management Solutions. In: Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services. SESSION: Privacy technologies, Markham, Ontario, Canada, vol. 380, Article No. 25 (2006)

15. Goodrich, M., Tamassia, R., Yao, D.: Notarized Federated Identity Management for Web Services. In: DBSec 2006, p. 133 (2006)
16. Khan, M.K., Zhang, J.: Multimodal Face and Fingerprint Biometrics Authentication on Space-Limited Tokens. Neurocomputing 71(13-15), 3026–3031 (2006)
17. Rundle, M., Ben Laurie, B.: Identity Management as a Cybersecurity Case Ctudy. In: Oxford Internet Institute Conference – Safety and Security in a Networked World: Balancing Cyber-Rights and Responsibilities, Research Publication No. 2006-01. The Berkman Center for Internet & Society Research Publication Series (September 2005), http://cyber.law.harvard.edu/publications
18. Madsen, P., Koga, Y., Takahashi, K.: Federated Identity Management for Protecting Users from ID Theft. In: Proceedings of the 2005 Workshop on Digital Identity Management, Fairfax, VA, USA, pp. 77–83 (2005)
19. Clauβ, S., Kesdogan, D., Klsch, T.: Privacy Enhancing Identity Management: Protection Against Re-Identification and Profiling. In: Proceedings of the 2005 workshop on Digital identity management. SESSION: DIM frameworks, Fairfax, VA, USA, pp. 84–93 (2005) ISBN:1-59593-232-1
20. Candebat, T., Gray, D.: Secure Pseudonym Management Using Mediated Identity-based Encryption. Journal of Computer Security 14(3), 249–267 (2006)
21. Gayathri, T., Venkadajothi, S., Kalaivani, S., Divya, C., Dhas, S., Sakunthala, R.: Mobile Multilayer IPsec Protocol. International Journal of Engineering and Technology 1(1), 23–29 (2009)
22. Wikipedia website, Identity Management Dimensions, http://www.en.wikipedia.org
23. Beres, Y., Baldwin, A., Casassa, M., Mont Shiu, S.: On Identity Assurance in the Presence of federated Identity Management Systems. Digital Identity Management, 27–35 (2007)

Secure Biometric Multi-Logon System Based on Current Authentication Technologies

Bobby L.Tait

University of Johannesburg
Kingsway South Africa
Btait@uj.ac.za

Abstract. The need for accurate authentication in the current IT world is of utmost importance. Users rely on current IT technologies to facilitate in day to day interactions with nearly all environments. Strong authentication technologies like the various biometric technologies have been in existence for many years. Many of these technologies, for instance fingerprint biometrics, have reached maturity. However, passwords and pins are still the most commonly used authentication mechanisms at this stage. An average user has to be authenticated in various situations during daily interaction with his or her environment, by means of a pin or a password. This results in many different passwords and pins that the user has to remember. The user will eventually either start documenting these passwords and pins, or often, simply use the same password and pin for all authentication situations.

This paper introduces a system developed to assist an average user with the host of passwords and pins that are needed for authentication. The system introduced will merge password and pin technologies with existing biometric technologies. These technologies will, in this system, interact with each other to provide optimum authentication.

Keywords: Security, Protection, Biometrics, Single sign on, Authentication, Identification.

1 Background

It is beyond doubt that all users need to be authenticated in many situations. Authentication according to Bruce Schneier [1] ensures that access control, integrity and confidentiality are being enforced. If the process of authentication fails, the access control mechanism fail, a message's confidentiality becomes suspect, and the integrity of transactions is in doubt.

It is universally accepted [2] that the information security services are defined as

- Identification & authentication
- Confidentiality
- Integrity
- Authorization
- Non-Repudiation

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 96–103, 2010.

The gate-keeper of all information security services is the service known as identification and authentication. We need to identify a user incontestably, and we need to be able to authenticate a person indisputably. If there is a way of subverting the first information security service of Identification and authentication, all the other information security services will fail.

Various mechanisms are at the user's disposal to assist with identification and authentication. However, at this stage passwords and pins are the most commonly used mechanism of authentication.

Users interact with various secured systems. For each secure system the user has a password. This password often has to conform to many rules (in terms of length, composition, age etc.)[16]. An average user might have five passwords or more to remember, subsequently, due to the variety of passwords that the user regularly needs, the user often documents all passwords, or uses a single- or hybrid password [3] for all systems.

A vast amount of research has demonstrated that passwords and pins are not secure, and can be subverted with little effort [4]. Many tools are available which will assist in the cracking and brute forcing of password files. Rainbow tables [15] are openly available to assist in faster cracking of large passwords (passwords consisting of 10 characters or more).

Biometrics on the other hand, is physically part of the user, thus, always with the user, and can under controlled situations successfully authenticate a user. Unfortunately, at this stage biometric technology is not as widely integrated into all systems as passwords are. Many factors preclude the common acceptance of a world-wide biometric system.

In the following sections, a system will be discussed that has been developed as part of a research project to seamlessly integrate biometrics into the current widely used password-based systems. It will allow the user to use his or her biometric identifier to gain access to all the areas where authentication is currently required by means of passwords and pins.

2 Biometric Technology

Biometrics: "(ancient Greek: bios ="life", metron ="measure") is the study of automated methods for uniquely recognizing humans based upon one or more intrinsic physical or behavioral traits" [6]. Ben Miller introduced the following definition in 1987 for biometrics: "Biometric technologies are automated methods of verifying or recognizing the identity of a living person based on a physical or behavioral characteristic" [7]. The international biometric industry association defines biometrics as "automated methods for verifying or identifying the identity of a living individual based on physiological or behavioral characteristics" [7].

Biometrics has the distinct advantage that it is always readily available, and in most cases, is difficult to steal from the user or to duplicate (unlike a password or pin).

The research combines current password-based authentication mechanisms with biometrics to allow a person to use any biometric technology as an authentication mechanism.

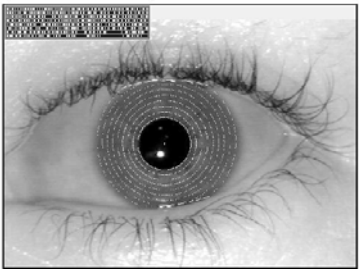


Fig. 1. Biometric Iris Recognition [5]

3 System Operational Environment

The operational environment of this system is defined as all the typical areas that a user will need to be authenticated. These environments include, but are not limited to:

- Email clients (programs or online web-based)
- Banking access pages
- Auction sites
- Public networking sites (Facebook, Twitter, Ning)
- Intranet access

Fig. 2. Example Logon (Windows Server 2003 R2)

It is clear from the above mentioned examples, that a user needs secure authentication in all cases. If authentication fails, a user is at risk of possible fraud or even identity theft to name only a few issues.

As indicated in figure 2, and also in each of the mentioned examples, for the purposes of identification and authentication, a user needs to supply a username and password. Very few systems currently use any other authentication mechanisms.

The next section introduces the mechanism of the biometric multi-logon system.

4 The Biometric Multi-Logon System

The system developed comprises of a relevant biometric reader and client software, which can be installed on a workstation or can be triggered from an USB flash drive, or any external storage medium.

The system is not limited to any particular biometric technology. For this reason a user can use the system in conjunction with fingerprint systems, iris systems or any mature affordable biometric system.

4.1 Overview of the System

The developed system acts as a mediator between current password- or pin based systems and a biometric solution. As pointed out, the particular biometric system used, is irrelevant, but it is important to note that certain biometric systems are inherently safer than other biometric systems.

In the case of the development conducted during the research, fingerprint technology was used, due to the affordability of the software development kit (SDK) and the optical sensor units.

To elucidate the mechanism of the system, a hypothetical situation is used of a user accessing a server by utilizing a remote desktop connection [9]. However, any situation requiring an username and password or pin, will operate seamlessly with this system.

As stated for this example, the user accesses the server by a remote desktop connection, the user is provided with a logon screen similar to the screen illustrated in figure 2, and illustrated in Diagram 1.

Step 1: The user opens a connection to a remote server. The authentication system of the remote server sends a challenge to the user requesting a user name for identification purposes, and a password for authentication purposes.

Step 2: The user responds, by supplying his biometric characteristic[1] to the biometric reader attached to his workstation, in this example, a fingerprint characteristic is supplied.

The biometric reader digitizes the provided biometric characteristic, and sends the resulting biometric data[2] to the user's workstation.

Diagram 1: Biometric Multi-Logon Overview

Step 3: The developed client software reacts to the interrupt from the biometric device's driver software, and tests the biometric data received against the reference biometric data[3]. This is done by using a decision algorithm to determine if the provided biometric data falls within the acceptable parameters for authentication. If the biometric data provided is accepted (and the user is thus now successfully authenticated), the system will continue to step 4. In all other cases, the user will be notified of the failure to authenticate.

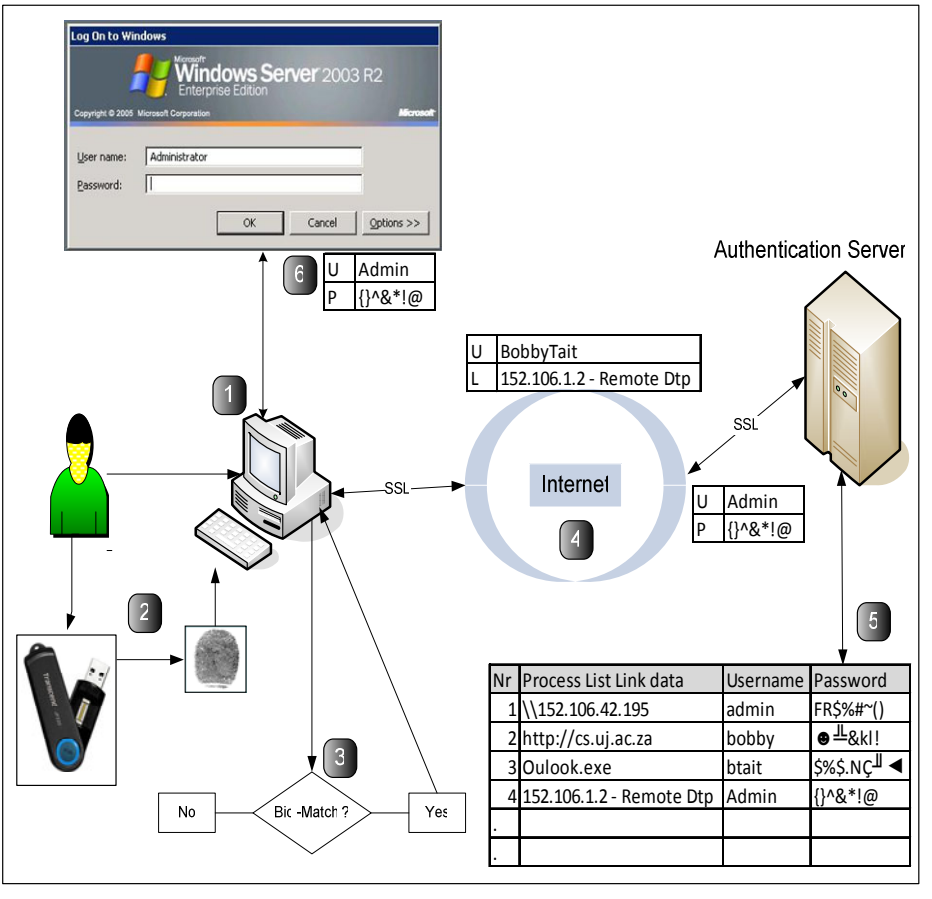
Step 4: Upon successful biometric authentication of the user in step 3, the client software consults an online authentication server by using a secure socket layer (SSL) [10] connection. The client software submits the user's identification information and the link information to the authentication server.

---

[1] The biometric characteristic that is physically part of the human, e.g. fingerprint, iris, retina.

[2] The digital representation of a biometric characteristic.

[3] Biometric data recorded during enrolment of the user, to facilitate in authentication of biometric data.



Step 5: The authentication server supports many users, all of which will have a unique ID. The server uses the unique ID received from the client software to ensure that the correct password repository is used.

In the illustrated example, the supplied link data is found in the user's password database, position nr 4. It shows that the user tries to access a remote desktop connection for the server found at the IP address 152.106.42.250.

The authentication server retrieves the user's username for the remote desktop connection at 152.106.42.250, as well as the password for this connection, as illustrated in diagram 1.

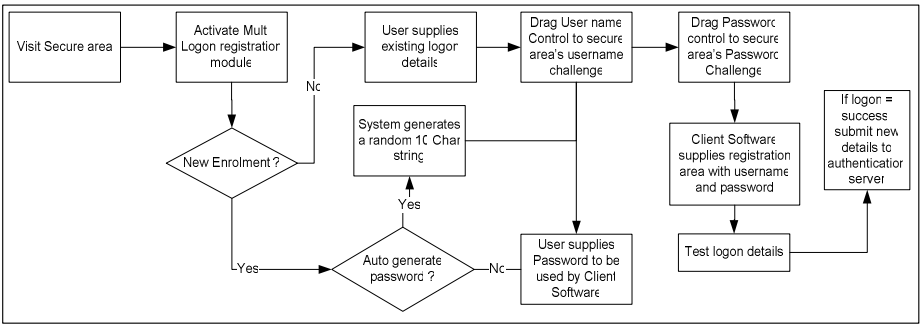
The authentication server sends the extracted user ID and password for the requested remote desktop connection, back to the client software over the established SSL connection.

Step 6: The client software receives the username and password for the remote desktop connection on IP address 152.106.42.250. It automatically populates the username name field of the logon challenge with "admin" (as per example) and the password field with "{}^&*!@", followed by a submit instruction.

These results in the user being logged on to the remote sever, without entering any password or username, simply by supplying a biometric characteristic.

4.2 Further Considerations

The whole system is automated; the enrolment process is important and for this reason will briefly be discussed.



4.3 Enrollment of a User

Diagram 2: Registration of a secure logon area

The biometric enrolment of the system is determined by the biometric characteristic that is used, and is not included in this paper. Literature describing the typical enrolment of a biometric characteristic is readily available [11], [12].

The focus of this section is to illustrate how a user will use the system to enroll a new secure area for biometric authentication. In this example as illustrated in diagram 2 the user links the multi-logon biometric system to his existing Facebook [13] profile.

Fig. 3. Facebook Logon Screen

Step 1: Visit the website or area that must be linked to the Multi-Logon biometric system. Considering that the user in this example is an existing user of Facebook, the user must supply the multi logon registration module with the username (email) and the password currently being used for Facebook.

Step 2: The user links the secure logon area (in this example for Facebook) with the registration module. This is accomplished by dragging the supplied username in the registration module to the username field of Facebook, and also dragging the supplied password in the registration module to the password field of Facebook.

Step 3: The registration module records the process in the task manager of Windows that is associated with the link that the user just made, and then automatically supplies the logon area of Facebook with the username and password that the user registered.

Step 4: The system tests the logon details supplied, to ensure that there is no error with the logon details before submission of the new details to the authentication server.

Step 5: if the logon for the newly registered area is successful, the client software submits the link details, username and password to the authentication server for future usage.

4.4 System Related Comments

If the user is not already registered on a secure logon area, the biometric multi-logon system will assist the user to add all the details related to normal registration automatically of a far as the system can recognize typical fields.

It is highly recommended that the user allows the biometric multi-logon system to generate a password, as the password generated will be at least 10 characters, which is absolutely random and mainly non-standard characters. Using long passwords, with non-standard characters, diminishes the possibility of a successful brute-force attack on the password-hash file significantly.

A user using the biometric multi-logon system, do not need to remember every password for all the secure areas anymore. The only action required from the user is to supply a biometric characteristic, and the system will log on to the secure area on behalf of the user, using the correct user ID and password.

5 Conclusion

The biometric multi-logon system integrates the functionality of a biometric system with the current password or pin based environments for identification and authentication. The user must in all cases simply supply a biometric characteristic, and the biometric multi-logon system will log the user on to various areas by using the user's passwords and usernames for the relevant areas. In essence the user does not need to remember any password or username ever again. The system is currently being tested by incorporating the client software onto a biometric-based USB flash drive [14]. By using such a flash drive, the user can use this system on any PC with internet access, as the biometric device is incorporated in the USB flash drive, and the client software is stored in the storage area of the flash drive.

References

[1] Schneier, B.: Secrets and Lies, digital security in a Networked world, ISBN 0-471-25311-1
[2] von Solms, S.H., Eloff, J.H., Smith, E.: Information Security, p. 8, ISBN 1-919774-39-4
[3] Stein, L.D.: Wen Security, A step-by-step reference guide, pp. 245–275, ISBN 0-201-63489-9
[4] LoptCrack security system, `http://www.l0phtcrack.com/`

[5] http://www.scan-r.eu/modules/pages/index.php?pagenum=6
[6] The International Biometric Industry Association, http://www.ibia.org
[7] Miller, B.: Biometric consortium Listserv (August 2, 2002)
[8] http://www.transcendusa.com/Products/ShowImg.asp?ModNo=169&vplay=yes
[9] Get started using Remote Desktop with Windows (July 25, 2006), http://www.microsoft.com/windowsxp/using/mobility/getstarted/remoteintro.mspx
[10] Whitman, M.E., Mattord, H.J.: Principles of Secure Information security, 3rd edn., pp. 382–383, 385, ISBN 978-0-8400-3116-7
[11] Woodward, J.D., Orleans, N.M., Higgins, P.: Biometrics: Identity assurance in the information age, pp. 31–33, ISBN 0-07-222227-1
[12] Liu, S., Silverman, M.: A Practical Guide to Biometric Security Technology. IT Pro, 27–32 (January/February 2001), IEEE 1520-9202/01
[13] Facebook community site, http://www.facebook.com
[14] Transend security series USB Flash Drives, http://www.transcendusa.com/Products/ModDetail.asp?ModNo=169&LangNo=0&Func1No=1&Func2No=76
[15] Free Rainbow Tables, http://www.freerainbowtables.com/
[16] Banff, A.: A large-scale study of web password habits. In: Proceedings of the 16th international conference on WWWeb, pp. 657–666, ISBN:978-1-59593-654-7

Analysis of Fingerprint Image to Verify a Person

Hossein Jahankhani and Maktuba Mohid

School of Computing, IT and Engineering
University of East London, UK
h.Jahankhani@uel.ac.uk

Abstract. Identification and authentication technologies are increasing day by day to protect people and goods from crime and terrorism. This paper is aimed to discuss fingerprint technology in depth and analysis of fingerprint image. Verify a person with a highlight on fingerprint matching. Some fingerprint matching algorithms are analysed and compared. The outcomes of the analysis has identified some major issues or factors of fingerprinting, which are location, rotation, clipping, noise, non-linear distortion sensitiveness/ insensitiveness properties, computational cost and accuracy level of fingerprint matching algorithms. Also a new fingerprint matching algorithm proposed in this research work. The proposed algorithm has used Euclidean distance, angle difference, type as matching parameters instead of specific location parameter (like, x or y coordinates), which makes the algorithm location and rotation insensitive. The matching of local neighbourhoods at each stage makes the algorithm non-linear distortion insensitive.

1 Introduction

Human bodies contain unique characteristics. Each of these characteristics differentiates one from another and therefore make easier to identify. The most traditional method for identification and authentication is the use of fingerprint technologies, [12]. This research project is bio-technology related project where biometric measurement is the main focus. The physiological biometrics such as fingerprints, iris and retinal scanners were developed in early stages. This research project based on fingerprinting process; especially fingerprint matching processes and related algorithms have been analysed, which leads to discover a new algorithm with some new features, [16] and [17].

A new algorithm has been proposed regarding those major factors where algorithmic solution has been given to overcome those major factors.

The target of this project is to find a standard, efficient and accurate fingerprint matching algorithm. To find efficient, standard and accurate fingerprint matching algorithm, first of all the factors/ issues by which the efficiency and accuracy of fingerprint matching algorithm can be measured need to be identified. The aim of this paper is to identified factors and issues by comparing and analysing different types of fingerprint matching algorithms. The factors/ issues are:

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 104–119, 2010.

1. Rotation insensitiveness
2. Location insensitiveness
3. Clipping insensitiveness
4. Non-linear distortion insensitiveness
5. Noise insensitiveness
6. Low cost
7. High accuracy level

2 Background

The word 'Bio' means biology and 'metric' means measure, so the combined word 'Biometric' means biological measurement. In general, biometric term is used relating to the identification of a person using biological characteristics with uniqueness property of human being. Two types of biological characteristics are used;

1. Behavioural – relating to the behaviour of a person. Like, voice recognition, typing rhythm etc. [14]
2. Physiological – relating to the shape, size or other properties of different parts of the body. It includes finger print recognition, face recognition, hand/ palm geometry recognition, DNA identification, retina-based measure etc.

All these types of biometric measures are almost unique for every person. Of them fingerprint recognition is the most popular one due to its immutability and uniqueness properties.

Immutability: Fingerprint pattern of a person does not change during his whole life other than the size and due to some special incidents, like accident, skin diseases, operation etc. This characteristic is called immutability.

Uniqueness: Fingerprint pattern is unique for every person, i.e. one person's fingerprint does not match with another person. This characteristic is called uniqueness. Even one person's fingerprints of two fingers do not match. Hand geometry, face recognition, behavioural biometric measures etc are not that much unique like fingerprint.

Also finger print image is easier to take than many other biometric measures. In case of retina-based measure, lighting is necessary.

Fingerprint patterns are normally categorized by friction ridges. The skin of inside surface of hands and fingers contain minute raised ridges known as 'friction ridge skin'. The minutiae of fingers can be categorized by some principal categories –

1. Ridge ending: the ending point of a ridge.
2. Bifurcation: The division of a ridge into two ridges.
3. Dot: Ridge like a dot, i.e. having equal width and length.
4. Spur: One type of bifurcation where a short ridge is a branch of a long ridge.
5. Lake/ Enclosure: One ridge divides into two ridges and then reunites again.
6. Short ridge: A ridge which is short in length.
7. Bridge: A bridge-ridge that connects two ridges.

Combinations of these different types of ridges form a pattern of fingerprint. The positions and numbers of these different types of ridges create different types of fingerprint patterns, which show uniqueness.

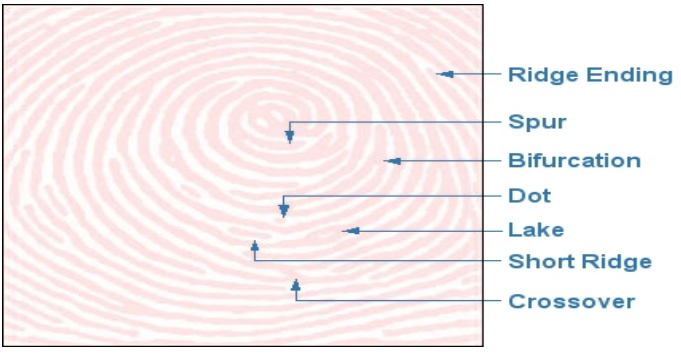


Fig. 1. The categorization of minutiae of fingerprints, Fingerprint analysis - the basics, [13]

Types of fingerprint patterns are available

1. Loop: Ridges enter from one side, make curves and then go out from the same side. The centre of a loop is called core. The area of triangulation and division of ridges (bifurcation) is called delta. These two focal points, i.e. core and delta make a loop. In case of loop pattern, core, delta, area of delta are recorded, [15].
2. Whorl: Circular shaped ridges. Two or more deltas make a whorl. In case of whorl pattern, deltas and areas of deltas are recorded.
3. Arch: Ridges enter from one side, make a raise and then go out from the opposite side. Arch does not contain any core or delta. In case of arch pattern, whole area is recorded.

The full process of fingerprint analysis is done using different steps:

1. Step 1: Image recording
2. Step 2: Image processing
3. Step 3: Image verification

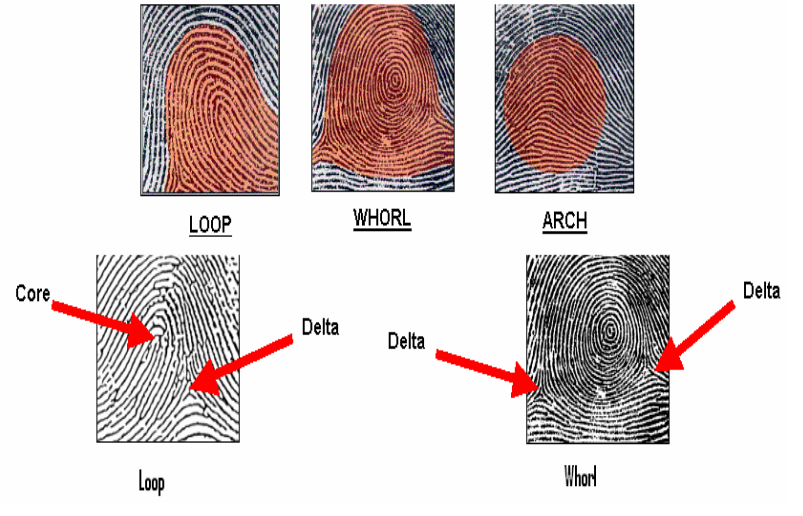


Fig. 2. Fingerprint patterns, legible fingerprints, Federal Bureau of Investigation

Different types of fingerprint recording impressions are used

Image recording is a very important part of fingerprinting. Image is taken over ink and paper, Porelon pad etc. Different types of devices or readers are used for recording fingerprint, like livescan devices, solid-state reader, optical reader, touch screen etc.

1. Rolled impression: In this case finger is rolled from one side of the finger nail to the other.
2. Plain impression: Finger is placed in 45^{o} angle and fixed over the recorded material. It is used to verify the accuracy of rolled impression of fingerprint.

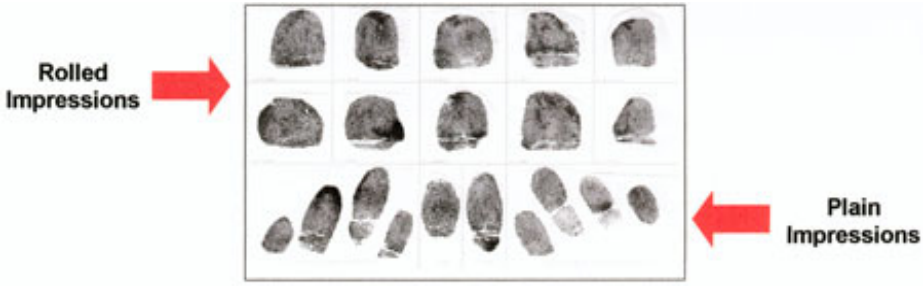


Fig. 3. Fingerprint recording, legible fingerprints, Federal Bureau of Investigation [11]

Image processing and image enhancement

Fingerprint is one of the most popular biometric measurement techniques. But the main problem arising regarding fingerprint is unclear image which is not suitable for image verification or might give wrong result. The image becomes unclear due to various reasons, like poor skin and ridge condition (skin disease, occupational mark etc), wrong way of taking fingerprint, wrong placement of fingers, poor quality of scanning device etc, [2], [4] and [5].

In this regards three types of error appears in fingerprint image:

False minutiae are created True minutiae are missed Position and orientation of minutiae are changed.

Investigations showed that three types of fingerprint image regions occur:

Well-defined region: In which region different categories of minutiae (ridges, bifurcation, spur etc) are easily recognized and extracted using minutiae extraction algorithm, is called well-defined region. Recoverable corrupted region: In which region different categories of minutiae are not easily recognized, but can be extracted with the help of neighbouring regions, is called recoverable corrupted region.

Unrecoverable corrupted region: In some regions of fingerprint image, different categories of minutiae are not totally recognized (sometimes invisible) due to noise or distortion. Even the neighbouring regions become unable to provide enough information to extract minutiae patterns from those corrupted regions. This type of corrupted region is called unrecoverable corrupted region.

So after getting fingerprint image, it needs to be processed or enhanced to make it clear and suitable for identification. Two types of enhancement methods are normally used:

Fig. 4. Poor quality fingerprint images, [1]

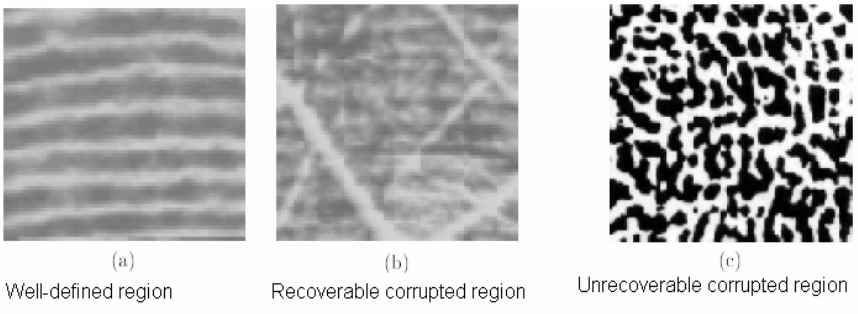


Fig. 5. Fingerprint regions, [1]

1) **Binarization-based method:** In this method, first of all image is transferred to binary image and then minutiae are extracted from the binary image.
2) **Direct gray-scale enhancement process:** Minutiae can be extracted from gray-scale image directly without transforming into binary image using a direct gray-scale enhancement algorithm.

Fingerprint matching

Fingerprint matching is one of the oldest, reliable and widely used biometric measures.

Normally two approaches are used in case of fingerprint matching:

1) Minutiae method: In minutiae method, the information of ending (terminal point) and bifurcation (separation point) of ridges, like type of ridges, position, direction etc are matched.
2) Pattern matching method: In pattern matching method, the input fingerprint image and registered fingerprint image are compared using direct comparison of flow of ridges of all places of the images.

Image transformation

Sometime enrolee and claimant fingerprints are not in a same format, like rotated in different angles, different scaling, different sizes etc. In those cases, fingerprint images need to be transformed to make the formats same. Otherwise, feature extraction, matching might be wrong, which might give wrong result. Different types of transformations are used:

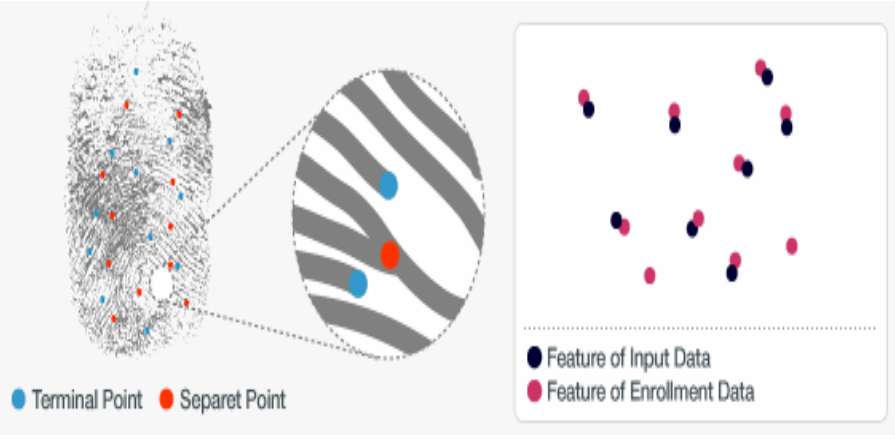


Fig. 6. Minutiae method, fingerprint Verification Algorithm, [10]

Fig. 7. Pattern matching method, fingerprint Verification Algorithm, [10]

1) Rigid transformation:
   Rotation, shifting etc are called rigid transformation.
2) Affine transformation: Scaling, shearing etc are called affine transformation.
3) Non-linear transformation:
   Sometimes some transformations need to be applied on the image with some complex equations. This type of transformation is applied using thin plate spline deformation model or by modelling as piece-wise linear transformation, [8], [18].

Fingerprint verification

In this case, first of all fingerprint of enrolee is recorded and put in the database with his/her identity. Later on, when anyone claims to be that person, the claimant fingerprint is matched with enrolee one for verification. It is used for accessing secured systems, entering secured areas etc. As one fingerprint is matched with another fingerprint in the database, it is called "one to one matching".

Different approaches or methods are used for fingerprint verification, which are –

1) Neighbourhood minutiae comparison
2) Modified approach of neighbourhood comparison.

3) Global matching/correlation approach
4) Modified correlation approaches

Acceptance/ rejection results, i.e. recognition rate, accuracy, as well as cost of fingerprint verification are dependent on the threshold value. There are two types of recognition rate: false acceptance rate (FAR)/ false match rate/ Type II error and false rejection rate (FRR)/ false non-match rate/ Type I error. If the threshold value is chosen to be very high, number of false rejections (FRR) will be increased. In this case, it might happen that fingerprints of same finger of same person, which should be matched, will not be matched due to noisy image. On the other side, if the threshold value is chosen to be very low, number of false acceptance (FAR) will be increased. In this case, it might match two fingerprints which should not be matched. So threshold value should be chosen in a range so that false rejection and false acceptance will not occur.

There are two types of ways of choosing threshold value:

1) Back-end adjustment
2) Front-end adjustment

Fingerprint identification

Fingerprint identification is used to identify criminals, where fingerprint left by criminal is matched with many fingerprints of database. As one fingerprint is matched with many unknown fingerprints, it is called "one to many matching".

In case of fingerprint identification, claimant fingerprint is matched with a large number (10 to ten million) of enrolee fingerprints. It will be time consuming if fingerprint verification process is applied to match each of these claimant-enrolee fingerprint pairs. That's why, in case of fingerprint identification, different approach is used, which is a two step process.

- **Step 1:** Test Fingerprint and fingerprints of database are categorized into different types of pattern classes (whorl, loop, arch etc). After classification process, pattern matching method is used for the comparison of test fingerprint and fingerprints of database. Step 1 is a fast process.
- **Step 2:** Minutiae method is applied over all of those pairs of claimant-enrolee fingerprints which show little difference in step 1 results. This step is time consuming.

Different types of fingerprint matching algorithms

There are different types of minutiae based fingerprint matching algorithms are available. Three of them have been discussed, analysed and compared here. They are graph-based fingerprint matching, fingerprint matching using onion layer algorithm of computational geometry and K-plot algorithm, [9].

Fingerprint matching using graph matching approach: The graph based fingerprint matching algorithm which has been discussed here is proposed by D.K. Isenor and S.G.Zaky (Department of Electrical Engineering, University of Toronto, Canada). The algorithm converts the whole image of fingerprint into graphs where ridges are considered to be nodes of graphs and relationships between ridges are considered to be the connections of nodes.

Following information is kept for each node/ ridge:

a) The ridge length;
b) Nature of minutia in each end (i.e. ending or bifurcation);
c) If any of the minutiae endings is bifurcation, the number of end neighbours involved in that end;
d) The information whether the ridge is complete or not;
e) The list of ridge neighbours ordering by the clockwise direction starting from end 1;
f) The overlap length for each neighbour.

After converting images into graphs, the graphs are matched using those information (given above) of ridges/ nodes. The noise reduction step makes the algorithm insensitive to noise, displacement and distortion.

Some steps are followed for graph based matching. The steps are given below:

- Step 1 – Ordering of sides and endings;
- Step 2 – Identifying neighbours;
- Step 3 – Numbering the levels;
- Step 4 – Drawing the graph;
- Step 5 – Reducing noise;
- Step 6 – Partitioning;
- Step 7 – Refinement;
- Step 8 – Scoring, [6].

Fingerprint matching using onion layer algorithm of computational geometry

This fingerprint matching approach has been proposed by [7]. The algorithm is based on onion layer algorithm of computational geometry. The algorithm generates nested polygons using repeated convex hull algorithm where nodes of the polygons are the minutiae points of fingerprint image. Thus it is minutiae based algorithm. The whole algorithm can be divided into three parts:

a) Finding nested polygons using convex hull algorithm and assigning depth number.
b) Finding reference triangle for calculating rigid transformation parameters and local matching
c) Pair matching (global matching) where it compares type, x, y coordinates and orientation angle of minutiae points.

This algorithm rejects unmatched fingerprints before global matching, which saves time by avoiding time consuming global matching steps.

Fingerprint matching using K-plet and Coupled BFS algorithm

This algorithm has been proposed by [3], (Centre for Unified Biometrics and Sensors, University at Buffalo, NY, USA). The algorithm is based on graph matching principles and is divided into 2 parts:

1. Graph representation named K-plet.
2. Matching using coupled BFS (Breadth First Search) algorithm.

A graph is drawn using K nearest neighbours of all minutiae points. After extracting minutiae points, this algorithm draws a graph using K nearest neighbours of all minutiae points where each neighbour contains following data --

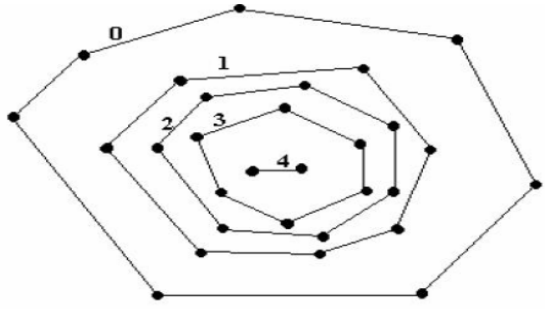


Fig. 8. Nested polygons constructed from point set, [7]

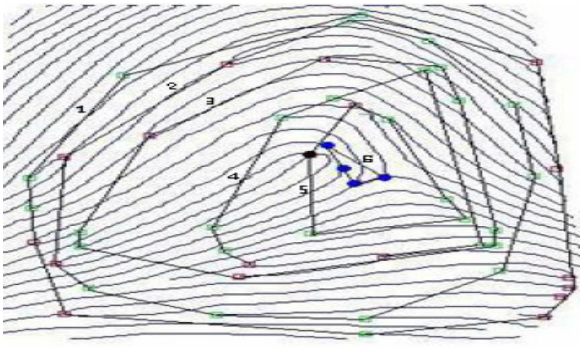


Fig. 9. Fingerprint with minutiae and its nested polygons, [7]

Euclidean distance, orientation angle difference, connecting edge direction between specific minutia point and the neighbour.

There are two approaches of finding K nearest neighbours:

1. K-nearest neighbours depending on Euclidian distances.
2. Nearest neighbour in each quadrant sequentially depending on Euclidian distances.

After drawing graphs, they are traversed and matched using coupled BFS (CBFS) algorithm. The CBFS algorithm is a dual graph traversing algorithm, i.e. it traverses both template and input fingerprint graphs at the same time and matches pair of minutiae points using their neighbour data. Two minutiae points are said to be matched only if all of their neighbours are matched.

As local neighbourhoods are matched at each stage, the algorithm is insensitive to non-linear distortion.

No explicit alignment is required in this algorithm as distance and angle differences are used instead of particular location parameters (like, x,y coordinates).

Experimental procedure

The algorithm which has been proposed in this thesis is minutiae based algorithm, which matches all the neighbours of all minutiae points.

The minutiae points are said to be matched if type and all neighbours are matched. Two neighbours of two minutiae points are said to be matched if type, Euclidian distance, angle difference are matched. The minutiae points and their neighbours are matched recursively.

The neighbours of any minutia point are sorted according to their Euclidian distances in increasing order. And one additional feature is added in case of matching neighbours of minutiae points, which decreases the cost of matching.

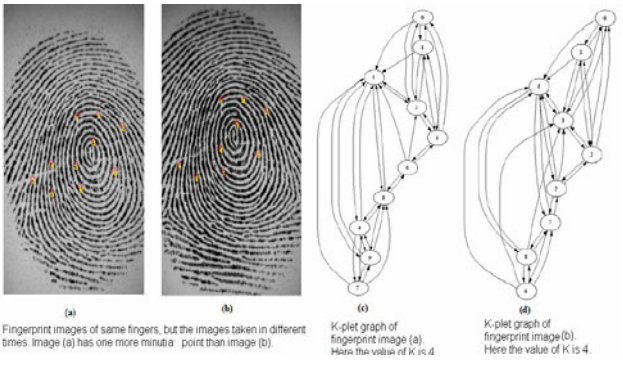


Fig. 10. K-plet fingerprint matching algorithm, [3]

The matching runs for all neighbours of all minutiae points of enrolee/template fingerprint against all neighbours of all minutiae points of claimant/input fingerprint. It stops matching if Euclidian distance of claimant/input neighbour becomes greater than Euclidian distance of enrolee/template neighbour. As the neighbours of any minutia point are sorted according to Euclidian distances in increasing order, if any Euclidian distance of any claimant/input neighbour becomes higher than Euclidian distance of enrolee/template neighbour, the Euclidian distances of later neighbours of claimant minutia point will be more higher than the current one and it will not match, so there is no point to continue matching, thus in this way it decreases some cost of matching. The feature has been clearly described in pseudo code and Matlab code.

Another modification has also been proposed, which will deal with the problem of clipping image. If any image is clipped from any side, the minutiae points of the discarded side will not be found and remain unmatched, which might affect the matching result. The proposed algorithm detects those minutiae points which remain in the discarded side and then it subtracts the numbers from the result, so that it might not affect the result.

At the time of matching the neighbours of two minutiae points, if any neighbour of any fingerprint image remains unmatched, the proposed algorithm investigates whether the x or y coordinate of that neighbour is greater or smaller than the x or y coordinates of all the minutiae points of another fingerprint image. That means it investigates whether the neighbour minutia point remains in the boundary of any side. Because clipping of image occurs in boundaries. If the point remains in the boundary, it then checks whether there is any other neighbour of the same image, which remains in the same side of boundary after that unmatched neighbour and which is matched. If any neighbour exists like that, the unmatched neighbour is not considered to be in the

discarded portion of cropped image. Otherwise, it is considered to be existed in the discarded portion of cropped image and then it is not considered or counted in calculation of matching score.

After doing matching, if any minutia point of any image remains unmatched, the algorithm does the same thing, i.e. investigates whether the minutia point remains in any side of the discarded portion of cropped image or not. If the minutia point remains in the discarded portion of cropped image, the algorithm does not count it in the calculation of final matching score. For matching purpose, this suggested algorithm use Euclidian distance, angle difference and type instead of using particular location parameters (like x, y coordinates), which makes the algorithm insensitive to location and rotation.

As local neighbourhoods are matched at each stage, the algorithm is insensitive to non-linear distortion as well.

All these features make this suggested algorithm insensitive to location, rotation, non-linear distortion and clipping. No explicit alignment is needed due to these insensitiveness properties, which decreases cost as a whole.

The algorithm has been described below

After extracting minutiae points, each minutia point has been assigned a unique ID and status. At first the statuses of all minutiae points have been set to unmatched. Then the neighbour data has been calculated for all minutiae points for both of the two images. The neighbour data contains neighbour minutia point ID, matching status, Euclidian distance and orientation angle difference between the neighbour and that minutia point. At first the matching statuses of all neighbours have been set to unmatched. The neighbours of any minutia point have been sorted according to their Euclidian distances in increasing order.

After gathering, calculating and collecting all data/information, the minutiae points of two images have been matched using Pair_Matching function (which will be described later) if the minutiae points do not have matched status value. After matching all minutiae points of two images, the total number of unmatched boundary points of two images, which remain in the discarded portion of cropped image, has been calculated using checkBoundaryPoint function (which will be discussed later). This function checks whether the unmatched point remains in the discarded portion of cropped image or not. If it remains in any side of discarded portion, it will not be counted in calculation of final result.

After this boundary point checking, the number of matched pairs has been calculated using statuses of minutiae points of enrolee image.

At last the matching score or result has been calculated using following equation:

$$\text{Result} = \frac{2 \times \text{Number of matched pairs of two fingerprints}}{\text{number of minutiae points of input fingerprint} + \text{number of minutiae points of template fingerprint} - \text{Total number of unmatched boundary minutiae points of two fingerprints, which remain in the discarded potion of clipped image}} \times 100$$

Here, 100 are multiplied to get a percentage value of the matching score. If the result is greater than or equal to the chosen threshold value, the two images will be said to be matched, otherwise, they will be said to be unmatched. The pseudo code and Matlab code of this function are described in appendix 3 and 4 correspondingly.

The Pair_Matching function is used for matching of two minutiae points of two fingerprints. The function proceeds only if status of none of the minutiae points is matched. In this function, a matching score has been calculated for the pair, which has been initially set to 0 values. Nested loops are present in this function. The outside loop is for the neighbours of enrolee/template minutia point and inner loop is for the neighbours of claimant/input minutia point. Inside two loops, two neighbours have been checked for matching. Two neighbours will be said to be matched if---

$$\textbf{Result} = \begin{cases} yes, if \quad |Dist_i - Dist_t| < r_0 \\ Type_i = Type_t \ and \\ |\theta_i - \theta_t| < \theta_0 \\ no, Otherwise \end{cases}$$

Here $Dist_i$ is the Euclidian distance between the neighbour and minutia point of input image, $Dist_t$ is the Euclidian distance between the neighbour and minutia point of template image, $Type_i$ is the type of the neighbour of input minutia point, $Type_t$ is the type of the neighbour of template minutia point, θ_i is the angle difference between the neighbour and minutia point of input image, θ_t is the angle difference between the neighbour and minutia point of template image. r_0 is the distance threshold, θ_0 is the angle threshold. That means two neighbours are said to be matched only if—

1. Their types are same;
2. Their Euclidean distance difference is lower than the distance threshold value;
3. The difference between their orientation angle differences is lower than the angle threshold value.

If the matching result comes yes, the matching statuses of both neighbours will become matched and matching score will be increased by one. If two neighbour minutiae points are matched and none of them has status (minutia point status) as matched, the Pair_Matching function will be called recursively using two neighbours as two minutiae points. If the Euclidian distance of current neighbour of input/claimant minutia point becomes higher than the threshold value plus the Euclidian distance of the current neighbour of template/enrolee minutia point, the code will exit from the inner loop, which saves cost of matching.

After completion of outer loop, all unmatched neighbours of both of the input and template minutiae points have been checked to be the boundary points of the discarded portion of cropped image using checkBoundaryPoint function.

After that, a matching result has been calculated for the minutia point pair using following equation:

$$\text{Result} = \frac{2 \times \text{Total matching score of the minutiae point pair}}{\text{number of neighbours of input minutia point} + \text{number of neighbours of template minutia point} - \text{total number of unmatched boundary neighbours of two minutiae points, which remain in the discarded portion of clipped image}} \times 100$$

Here, a value 100 is multiplied to get a percentage of the result. If the result becomes greater or equal to the threshold value, two minutiae points will be said to be matched, otherwise, they will be said to be unmatched. If two minutiae points are matched, matched as status values will be assigned to them. checkBoundaryPoint function is used to check whether any unmatched minutia point of any fingerprint image exists in the discarded portion of another clipped fingerprint image. It checks whether x or y coordinate of the specific minutia point is greater or smaller than corresponding x or y coordinates of all minutiae points of another image. If x or y coordinate of the point is greater or smaller than corresponding x or y coordinates of all minutiae points of another image and if there exists no matched minutia point in the same portion of the same image after that unmatched minutia point, it will consider the point to be placed in the discarded part of the clipped image, and will get the reason of being unmatched.

The checkBoundaryPoint function works for each unmatched point of each image at a time. Nine steps are used in this case where eight steps are used for checking. The steps are:

1. Checks whether the x coordinate of the specific point of that image is greater than the x coordinates of all minutiae points of the counter image.
2. If the result of step 1 is true, it checks whether there exists any minutia point in the same image, whose status is matched and the x coordinate is greater or equals to the x coordinate of the specific point. If any point is found in this step, it passes to step 3, otherwise passes to step 9.
3. Checks whether the x coordinate of the specific point of that image is smaller than the x coordinates of all minutiae points of the counter image.
4. If the result of step 3 is true, it checks whether there exists any minutia point in the same image, whose status is matched and the x coordinate is smaller or equals to the x coordinate of the specific point. If any point is found it passes to step 5, otherwise passes to step 9.
5. Checks whether the y coordinate of the specific point of that image is greater than the y coordinates of all minutiae points of the counter image.
6. If the result of step 5 is true, it checks whether there exists any minutia point in the same image, whose status is matched and the y coordinate is greater or equals to the y coordinate of the specific point. If any point is found in this step, it passes to step 7, otherwise passes to step 9.
7. Checks whether the y coordinate of the specific point of that image is smaller than the y coordinates of all minutiae points of the counter image.
8. If the result of step 7 is true, it checks whether there exists any minutia point in the same image, whose status is matched and the y coordinate is smaller or equals to the y coordinate of the specific point. If any point is not found in this step, it passes to step 9

9. It considers the specific point to be existed in the discarded portion of clipped image. This boundary point checking is done both for unmatched neighbours and unmatched minutiae points of both template and input fingerprint images. This function solves the problem of clipped/cropped image.

Cost

The full matching algorithm is divided into 3 parts:

a) Sorting of neighbours for all minutiae points of two images
b) Recursive BFS algorithm for pair matching
c) Boundary point checking for unmatched neighbours and minutiae points of both of the images

Some sorting algorithms take linear time to run, i.e. cost is O(M), where M is the number of minutiae points. If the sorting of neighbours is done for all the minutiae points, the cost will be $O(M^2)$. As it is done for two fingerprint images, total cost will be $O(2M^2)=O(M^2)$.

The cost of BFS (Breadth first search) algorithm is O(|M|+|N|), where M is the number of minutiae points and N is the number of neighbours. As all other minutiae points are considered to be the neighbour, |N|=|M|-1. Thus the cost of CBFS/ recursive pair matching algorithm will be O(|M|+|M|-1)=O(M). As neighbours are sorted according to distances and loop does not continue matching if the distance of enrolee/template neighbour becomes less than that of claimant/input neighbour, the cost will be less than O(M). The boundary point checking is done in at most eight checking steps. So the cost of the boundary point checking of every unmatched neighbour or minutia point is O(8M)=O(M), where M is the number of minutiae points. If number of unmatched neighbours or minutiae points is C, the total cost will be O(CM)=O(M). So the total cost of matching will be $O(M^2)+O(M)+O(M)=O(M^2)$.

Analysis

The three matching algorithms have been discussed, which are graph-based algorithm, K-plet algorithm, onion layer algorithm using computational geometry. Two algorithms from the three have been implemented, which are K-plet algorithm and computational geometry-based algorithm. After implementation, the three algorithms are analysed and compared.

3 Conclusion

The main focus of this research work was fingerprint matching. There are different types of fingerprint matching processes or algorithms. Of them, some are discussed, implemented, analysed and compared. The main target of this thesis is to find an algorithm which increases performance and accuracy levels. After investigating a lot, the following key points have been gathered which deal with the accuracy and performance issues of any fingerprint matching process:

Elastic distortion, location, rotation, clipping, noise insensitiveness and running cost. The computational geometry-based algorithm is non-linear distortion, location, rotation, clipping sensitive. For these sensitiveness properties, this algorithm needs to generate reference triangle to get transformation parameters, which increases the cost

of the algorithm. Also, some data (type, orientation angle of minutia point and edges) is stored twice, which also increases storage cost.

The graph-based approach is noise insensitive. But for reducing noise, it uses the refinement process, which costs higher. The algorithm is rotation and location insensitive, but it is clipping and non-linear distortion sensitive.

The K-plet algorithm is non-linear distortion, location, rotation insensitive, but it is clipping sensitive. Furthermore chosen value of K affects the cost and accuracy of the algorithm. If higher value of K is chosen, the accuracy level becomes high, but cost will increase as well. On the other side, if lower value of K is chosen, the cost becomes low, but accuracy level also becomes low.

Thus, after the investigation, in the conclusion, it can be said that the following properties should be highlighted while choosing/selecting any fingerprint matching algorithm:

- Rotation insensitiveness
- Location insensitiveness
- Clipping insensitiveness
- Non-linear distortion insensitiveness
- Noise insensitiveness
- Low cost
- High accuracy level

That means chosen algorithm needs to have high accuracy level and low cost.

For increasing accuracy level, either the algorithm needs to be rotation, location, clipping, non-linear distortion insensitive or needs some transformation processes. Noise is another factor which deals with accuracy issue. Noise reduction is highly dependent on image enhancement processes of image enhancement stage. Noise can also be reduced by some refinement processes while matching.

References

[1] Hong, L., Wan, Y., Jain, A.: Fingerprint image enhancement: algorithm and performance evaluation 20(8), 777–789 (1998), http://www.cs.ru.ac.za/courses/Honours/mmcourse/security/biometrics/MSU-CPS-97-35.pdf (Accessed July 4, 2009)

[2] Greenberg, S., Aladjem, M., Kogan, D.: Fingerprint image enhancement using filtering techniques 8(3), 227–236 (2002), http://www.cs.kent.edu/~wcheng/Presentation%201%20Image%20Enhancement/Fingerprint%20Image%20Enhancement%20using%20FilterTechniques.pdf (July 5, 2009)

[3] Chikkerur, S., Cartwright, A.N., Govindaraju, V.: K-plet and CBFS: A Graph Based Fingerprint Representation and Matching Algorithm (2006), http://web.mit.edu/sharat/www/research/papers/kplet_long.pdf (July 21, 2009)

[4] Kaur, M., Singh, M., Girdhar, A., Sandhu, P.S.: Fingerprint Verification System using Minutiae Extraction Technique (2008), http://www.waset.org/journals/waset/v46/v46-85.pdf (Accessed July 23, 2009)

[5] O'Gorman, L.: Fingerprint Verification (479), 43–64 (1999), http://www.cse.msu.edu/~cse891/Sect601/textbook/2.pdf (Accessed July 25, 2009)
[6] Isenor, D.K., Zaky, S.G.: Fingerprint Identification using Graph Matching* 19(2), 113–122 (1986)
[7] Khazaei, H., Mohades, A.: Fingerprint Matching and Classification using an Onion Layer algorithm of Computational Geometry 1(1) (2007), http://www.naun.org/journals/mcs/mcs-5.pdf (Accessed July 26, 2009)
[8] Chikkerur, S., Wu, C., Govindaraju, V.A.: Systematic Approach for Feature Extraction in Fingerprint Images, vol. 3072, pp. 344–350 (2004), http://www.cubs.buffalo.edu/pdf/finger-systematic.pdf (September 15, 2009)
[9] Corman, T.H., Leiserson, C.E., Rivest, R.L.: Introduction to Algorithms. MIT Press, Cambridge (1990)
[10] DDS (Digital development System): Fingerprint Verification Algorithm (2006), http://www.dds.co.jp/en/fv/algorithm.html (Accessed July 21, 2009)
[11] Federal Bureau of Investigation: Taking legible fingerprints, http://www.fbi.gov/hq/cjisd/takingfps.html (Accessed June 20, 2009)
[12] Grannblog: Problems with fingerprints for authentication, http://blog.granneman.com/2006/05/13/problems-with-fingerprints-for-authentication/ (Accessed June 20, 2009)
[13] Crimtrac: Fingerprint analysis – the basics (2008), http://www.crimtrac.gov.au/systems_projects/FingerprintAnalysis-TheBasics.html (Accessed June 20, 2009)
[14] IT University Copenhagen: Morphology, http://www.itu.dk/courses/MBSB/E2005/Download/morphology.pdf (July 5, 2009)
[15] Atalasoft DotImage 8.0: Morphology (2009), http://www.atalasoft.com/Products/dotimage/docs/Morphology.html (July 5, 2009)
[16] University of Wales Aberystwyth: Aberystwyth Users site (2007), http://users.aber.ac.uk/sac4/cs180/index.html (Accessed June 20, 2009)
[17] Wikipedia: Biometrics (2009), http://en.wikipedia.org/wiki/Biometrics (Accessed June 20, 2009)
[18] Wikipedia: Transformation matrix (2009), http://en.wikipedia.org/wiki/Transformation_matrix (Accessed July 22, 2009)

Methods of Organizational Information Security
(A Literature Review)

José Martins[1] and Henrique dos Santos[2]

[1] Academia Militar - Cinamil, Lisboa, Portugal
[2] University of Minho - Department of Information Systems, Guimarães, Portugal
jose.carloslm@gmail.com, hsantos@dsi.uminho.pt

Abstract. The principle objective of this article is to present a literature review for the methods used in the security of information at the level of organizations. Some of the principle problems are identified and a first group of relevant dimensions is presented for an efficient management of information security. The study is based on the literature review made, using some of the more relevant certified articles of this theme, in international reports and in the principle norms of management of information security. From the readings that were done, we identified some of the methods oriented for risk management, norms of certification and good practice of security of information. Some of the norms are oriented for the certification of the product or system and others oriented to the processes of the business. There are also studies with the proposal of Frameworks that suggest the integration of different approaches with the foundation of norms focused on technologies, in processes and taking into consideration the organizational and human environment of the organizations. In our perspective, the biggest contribute to the security of information is the development of a method of security of information for an organization in a conflicting environment. This should make available the security of information, against the possible dimensions of attack that the threats could exploit, through the vulnerability of the organizational actives. This method should support the new concepts of "*Network centric warfare*", "*Information superiority*" and "*Information warfare*" especially developed in this last decade, where information is seen simultaneously as a weapon and as a target.

Keywords: Method of Security of Information, Information Security, Dimensions of Security of Information, Information Security Management and Threats.

1 Introduction

The principle objective of this article is to present a literature review regarding methods of the security of information at the level of organizations, searching to identify some of the principle problems and also present a first group of relevant dimensions for an efficient management of the security of information.

We search to present a holistic vision of the security of information, according to its relevance and especially for the deciders of the military organizations. Guaranteeing that one understands and consequently searches to develop the organizational

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 120–130, 2010.

culture of the security of information to the collaborators at the various organizational levels (i.e. strategic, tactic, and operational) is of great importance.

In the majority of organizations and in the Military in particular, information is one of the most important active. This supports all the processes of the business, having to guarantee permanently the fundamental properties of the security of information: confidentiality, integrity, and availability [1-3].

Information is exposed fundamentally to three elements: technology, as the component that permits to save, process and transmit the information; people, or rather all the Stakeholders that could access the information through private networks or through the Internet and the negotiation process utilized in the manipulation of information [1].

To minimize the security of information risks associated to the mentioned elements, one demands fundamentally the correct identification of threats and the analysis of the vulnerabilities of the organizational actives that support the fluxes of information [3]. One can this way, realistically, put into action the simulation of possible attacks in which the information resource is subject to, in a way of determining the impact to the organization of an eventual attack.

With the identification of possible threats to the security of information, we could use International reports [4], obtaining a perception of the possible risks of the security of information in which the organizations are subject to. Some of the identified threats, could use a group of methods of attack to exploit one or more vulnerabilities of the organizational actives.

The possible methods of attack in which the Information Systems (IS) can be subject to, are in our study framed within the Information Operations. These consist principally of a group of activities and capacities utilized to affect the opponent's information and of its IS [5-6]. We could refer to some of the possible methods of attacks focusing on technology, for example: Malware, Denial of service, Packet Sniffer, Masquerade and the Man-in-the-middle [7].

In the context of Information Warfare, in which the Operations of Information could be utilized, these actions are developed to obtain the superiority of information, which consists of obtaining an operational advantage as result of the capacity of gathering, processing and disseminating a flux of information while one explores or denies to the opponent this same capacity [8-9].

Being that Information Warfare can be defined as a group of actions destined to preserve our Information Systems from exploitation, corruption, and destruction, while simultaneously one exploits, corrupts, or destroys the Information Systems of an opponent [8].

In the mean time other possible methods of attack exist, as those that are supported by attacks of Social Engineering that search to fundamentally manipulate the human element. Also, the more traditional that search for the physical destruction of the support to the infrastructure of the organization, should be a worry for the deciders. [10]

The effects of some of these methods of attack, of technical character, could be identified in the Cyber attack launched against Estonia in April and May of 2007, during which all activities of the Government were paralyzed [11], or in the Georgian conflict [12]. The conditions of "*Cyber Warfare and Computer Network Exploitation*" had outlined, by part of some of the countries such as China [13], potentially the relevance of organizational information security management.

It becomes clear with the literature review, that some of the issues regarding Information Security are solved solely with technical situations (e.g. Security of Communication). Others are so complex that only the purely technical path is not the solution, as in the case of security information management. A holistic vision is necessary in comparison to the various origins of risks of information security [1], being these internal or external to the organization and it natural, technological or human causes.

To make the situation even better, the appearance of new concepts came to be, developed especially in this last decade, as in the case of *Network centric warfare*" [14], "*Information superiority*" [15] and "*Information warfare*" [8].

In these new concepts, information has been seen simultaneously as a weapon and as a target [16], which brings us to the necessity of new approaches in the security of information, being this possibly depicted as a "*Defensive Battle*" [17].

This article is divided in three sections, reflecting an approach focused on the revision of realized literature. In the first section, the framework of the problem is depicted. In the second section, the principle approaches to the security of organizational information are presented and simultaneously, we present some of the possible dimensions to be considered with information security management. The conclusions of the three sections are also presented.

2 Methods of Information Security

The following are the readings completed: ISI Web of Knowledge, Google Scholar, and Scopus, with the following key words/expressions: Information Security Standards, Information Security Management and Frameworks of Information Security.

Simultaneously the following academic books of reference were taken into consideration Computer Science and Information Systems Technologies that approached some of the principle topics regarding information security with focus on technology [7, 18-19].

We search to also, obtain the principle referential of knowledge of the professionals of security of information. As such, the following was indicated in detail one of the professional certifications of information security systems of most known by the industry, the Certified Information Systems Security Professional (CISSP) [20].

2.1 State Of The Art

The principle objective of information security is to avoid that the threats explore the vulnerabilities of the organizational actives. Considering the threat, as the potential cause of an incident, in which damage could result in a system or in the organization and that the understood vulnerability as the weakness of an active or group of actives, that could be explored by one or more threats [21].

These threats utilized methods of attack, in a way of exploring the vulnerabilities of the actives of the organization and this way causing impact. The methods of attack become reality to the group of action utilized by a threat (that has potential) to explore one or more vulnerabilities of a determined active.

In the organizations information security has as a principle objective guaranteeing confidentiality, integrity and the availability of information. With the confidentiality

one searches to guarantee that the information is accessible to only those that are authorized to have access to them. In the case of availability, the objective is to guarantee that the authorized users have access to the information when necessary. Finally with integrity, one searches to guarantee that the information and or ones methods of process are not modified in an abrupt manner [3].

The contributes of investigation in information security are centered according to the readings done by Siponen & Oinas-Kukkonen (2007), fundamentally in four areas of investigation: Access to Information Systems, Secure Communication, Security Management and Development of Secure Information Systems. Of these issues, one should focus their reading in security management, more specifically in information security management.

Information is one of the most important actives for the organization [22], weather they have lucrative ends or not. Being it fundamental to ones security, through the utilization of controls. These could consist in a policy, procedure, practice, or organizational structure designed in a manner of proportioning a reasonable degree of certainty of what unwanted occurrences are prevented, highlighted, or corrected [23].

Information Security Management is a process of structured management implemented in the organization, to guarantee the principle requirements of information security (confidentiality, integrity and availability); especially, information that is critical to support the negotiation processes of the organization. Due to the complexity of its management, it is necessary that it is supported through a technological tool. It being also fundamental that aside from the protection of information, it be guaranteed the protection of technological resources that support the fluxes of information [24].

The International norm ISO/IEC 27001 considers that information security management, as a process of structured management that permits to guarantee the principle requirements of information security, providing a model to establish, implement, operate, monitor, review, maintain and improve the Information Security System of Management (ISSM). Simultaneously it underlines the adoption of an approach by processes, which encompasses the utilization of a model of process that allows one to plan, execute, verify, and act upon all the processes of ISSM. This is also known as the PDCA (Plan, Do, Check, and Act) model.

Guaranteeing information security it fundamental to the security of the Information Systems, it is these that allow for the creation, transfer, and storage of the information throughout the organizational negotiation processes.

According to Baskerville (1993) the methods of analysis and the sketch of security of the information systems, are of three generations: Checklist Methods, Mechanistic Engineering Methods and Logical-Transformational Methods. These could be utilized as possible approaches for the security of information of organizations.

It is never the less custom in the organizations to initialize the process of information security with the identification and evaluation of one's risk [25]. Where the risk in information security is the possibility of a threat exploring its vulnerabilities of an active or a group of actives, in which could result in damages for the organizations. The risk is measured in terms of the combination of probability of an event occurring (e.g. a threat exploring vulnerabilities) and the losses or damages caused in an active or group of actives from the organization [21].

Whether it is the method of utilization, it is necessary that the information security be sent to the three levels of organization management, or rather to the strategic level, tacit and operational [26]. The executives accept this as a component of its "*corporate*

governance" [1], fundamental condition to implement a culture of security in all of the organization.

Simultaneously, for a correct understanding of information security it is necessary to have a holistic vision of the diverse aspects, as in the title of the example, the technological aspects, those related with the process of negotiation, the organizational and the individual behavior [27-28].

It is utile to put into practice the norms that permit for the integration of information management security, as the management of the business, using models of recognized management [28]. There are scientific studies that relate to the good practices of information security referred to in norms and used in companies with objectives of information security or rather with confidentiality, integrity and availability [29].

Due to the diverse studies identified and analyzed, with focus on security management of information at the organizational level, there are some difficulties in the application of the actual methods, in which we outline the following:

- In first case, the difficulty in calculating the probability of a threat exploring the vulnerabilities of an active [30], by evaluating the value of an active and consequently proving that the investment in information security has an adequate return [25];
- Also the multidimensional aspect of the security problem of the IS (ex. Threats, vulnerabilities, actives and impact) [30] increase the difficulty of the problem and of its possible solution;
- The nonexistence of a theory that covers all the information security management, through the integration of the diverse support theories [22];
- Fundamentally, and finally, the information security management in the organization, seems to not have reached a level of maturity that could turn into a repeatable management process [26].

In the realm of IS Security and of information, there could be diverse norms being utilized, codes of good practices, certifications, methods and methodologies considering one's specifications [31].

There does not exist, at present a group of knowledge for the information security accepted by the academic community in general, by the professionals of information security and by the industry, despite the fact that the intention exists [32].

2.2 Approaches to Information Security

At present, from the various norms, codes of good practices, guidelines, methods and certifications at a national, international or organizational level, to guarantee information security, we could find approaches that are more focused on technologies or on the negation processes. [31].

One of the principle International norms utilized in the management of information security is ISO/IEC 27001:2005 [28, 33-34]. There is no systematic method of development for its organizational application that could become a process for repeated management.

This norm presents the necessary requirements to implement a Management System of Information Security, being able to simultaneously be utilized to certify an organization in the management of information security defining its area of certification. It is

supported by the norm ISO/IEC 27002:2007, as a code of good practices for the management of information security, that posses a group of security controls with foundation on the good practices of the present security [35].

At the technological level, there are guidelines of security for the organizations, in a way of maintaining secure the information technological environment. We could refer as most relevant the group of norms at the international level that are part of the ISO/IEC 13335. Some of the principle technological controls are presented and considered to confront the most relevant attacks about the IS [36] and which focus on the controls of technology to be applied to the security of networks and communications [37]. Allowing through its utilization, planning and putting into practice the management of the most relevant aspects of operational security in an environment of information technology [38].

In the United States of America the recommendations of the National Institute of Standards and Technology, are the principle reference of orientation for the security of information systems supported by technology [39] and to effect tests in information networks in a way of identifying its vulnerabilities [40], and others. The objective of these publications it to provide directives for the selection and specialization of security controls for the organizations information systems.

There also are norms of technical security, like the ISO/IEC 15408:2005, that permits the evaluation of the level of security for the Information Technologies [41]**.** It is destined essentially to defining a group of criteria that permits evaluating the security of a determined product (e.g. *Windows Seven*) or a system (i.e. Oracle running over the Operational System Unix) and consequently allow for a comparison between different products or systems in experiment [42].

This norm is an adaption of the Common Criteria for Information Technology Security Evaluation (abbreviated as Common Criteria or CC)**.** Being that the Common Criteria resulted by the integration of Trusted Computer Security Evaluation Criteria (TCSEC), further known as "*Orange Book"* and the Information Technology Security Evaluation Criteria (ITSEC).

At the level of methods, we could refer the OCTAVE (i.e. Operationally Critical Threat, Asset, and Vulnerability) for evaluation and implementation of security in an organization, as it allows us to identify the actives that are important for its mission, the threats to the actives and the vulnerabilities that are exposed to the threats. This method consists of a group of sessions where the collaborators that work in this analyzed area of the organization define the risks and the methods of protection to be implemented [43].

We could also find studies, where there is a search to sort information security with the objective of the negotiation. Searching to construct a Balanced Scorecard with the indicators (e.g. metrics) possible of measuring the contribute of the investment realized in information security for the business of the organization [44].

We analyzed in detail one of the Professional certifications of information system security most recognized by the industry, the "*Certified Information Systems Security Professional*" (CISSP). We verified that there exists a group of fundamental dimensions to guarantee the security of information. We verified that there exists a group of fundamental dimensions to guarantee the security of information and consequently the fluxes of information.

There exists in this certification, the integration of topics of security of technological areas, as for example cryptology, the security of information networks and

telecommunications and the correct development of software and the control of access (i.e. identification, authentication, authorization, and responsibility). It is also fundamentally the management of risk of the security of information and the existence of a plan of continuation of the business and of Disaster Recovery. The physical security and the human factors that contribute to the security of information are referred, taking into consideration the threats that could affect the organization. Consequently, it is necessary the knowledge of the organization, its business processes, the laws and regulations that govern its activities [20].

However, the themes of CISSP are presented separately, without there being a method that permits its application as a process of divisible management in the organizations, making it possible for the integration of the diverse controls of information security presented in the themes. In military terms, NATO (North Atlantic Treaty Organization) defined a model that could be depicted in Figure 1, where the fundamental elements of security are related. This model separates the worried related with the communications and the computers, defining superiorly, a group of nodes that separate the objects of analysis. Also in this particular case there doesn't exist a method that allows one to integrate all these elements (eventually by our lack of knowledge due to the classification of information).

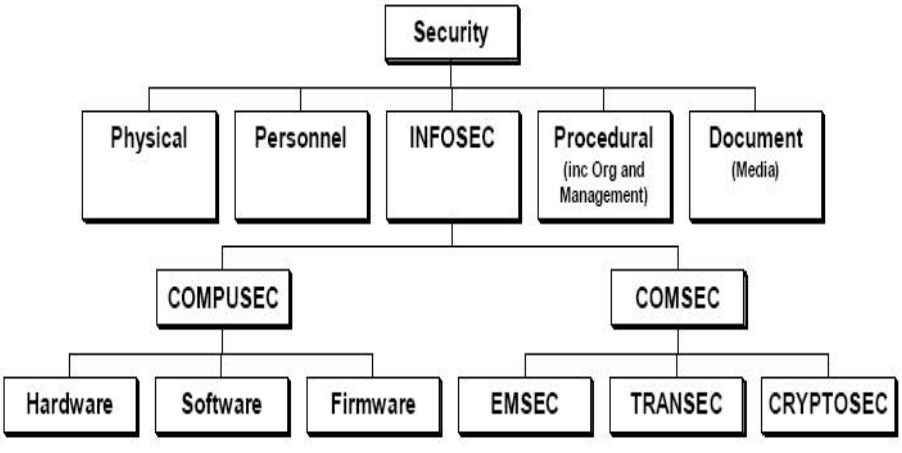


Fig. 1. NATO Model Of Security, Source: NATO [1]

From the literature review, we gathered the perception of some possible dimensions of information security to consider in the application of the method and according with the analysis of international norms, identify the dimensions of the organizational information security, the physical, the human, and the technological, following the different dimensions according to the indicated studies [45].

In conclusion, from the gathered information, we encounter methods oriented for the management of risk (e.g. OCTAVE). Being it understood that the management of risk of

[1] NATO restricted Web Site (version 1.6, 8 October 2007).

Information Systems, as the process of identifying, controlling, minimizing or eliminating the risks of security that could affect the information systems, at an acceptable cost [3].

We could identify norms of certification and of good practices of information security (e.g. ISO/IEC 27001 and ISO/IEC 27002). It also referred to norms oriented to the certification of the product or system (e.g. ISO/IEC 15408); Guidelines (e.g. ISO/IEC 13335-4, NIST 800-54) and some of the norms oriented more towards the negotiation processes (e.g. CobiT).

There also exists studies with proposals of Frameworks that suggest the integration of different approaches [46], with foundations on norms focused on technologies, in processes and taking into consideration the organizational and human environment of the organizations.

In essence, all of them permit information security, applied individually or in group, while from the completed literature review it was not possible to identify any method, which focuses on information security, in a conflicting environment.

The application of the method should be centered on the concept of Information warfare, given the specificity of the military organization. Consequently, the greatest contribution to the investigation of information security is the development of the method that permits integrating all the probable dimensions of information security, in a way of reducing to the minimum the risks of information security.

3 Conclusion

In this article the "*State of the Art*" is presented according to the utilized methods of information security at the level of the organizations, taking into consideration that in this context there exists a necessity of having a holistic vision of the dimensions of information security, especially within this last decade, where the information is seen simultaneously as an arm and target, in result of the new concepts of *Network centric warfare*", "*Information superiority*" and "*Information warfare*" developed, especially in the Military surroundings.

The study is based on the literature review, utilizing some of the most relevant scientific article of this area, in international reports and on the principle norms of management of information security. This way we can identify methods oriented for the management of risk, norms of certification and good practices of information security, norms oriented to the certification of the product or system and others oriented to the business process.

There also exists as proposals of Frameworks that suggest the integration of different approaches with grounds on the focused norms in technologies, in processes and taking in consideration the organizational and human environment of the organizations. Furthermore it was not possible to identify a method, which focus was on information security for military organizations integrated in a conflicting environment.

As possible themes of research that remain open, we could refer to:

- The development of a body of knowledge (theory) for the security of information accepted by all the academic communities, by the professionals of information security and by the Industry.

- The construction, the evaluation, the theorization and finally the justification in a methodology that permits the integration as the models of business in the organizations. Allowing for them to be supported in a Balanced Scorecard with possible indicators of measuring the contribution of realized investment in information security for the business of the organization.
- The development of Ontology of support to the principle concepts of information security. Due to the specifications of the military organizations permit for the construction of scenarios of virtual wars. Contributing in this way for the military decision made with the application of security controls, as if a battle field were to be referred to here.

We consider that the information security is a process of management and not a technological process, in which there exists a balance of powers between the most relevant dimensions of information security.

References

1. Posthumus, S., Von Solms, R.: A framework for the governance of information security. Computers & Security 23(8), 638–646 (2004)
2. Siponen, M., Oinas-Kukkonen, H.: A review of information security issues and respective research contributions. ACM SIGMIS Database 38(1), 80 (2007)
3. ISO/IEC27001: Information technology – Security techniques – Information Security Management Systems - Requirements (2005)
4. Richardson, R.: The 13th Annual Computer Crime and Security Survey, Computer Security Institute (2008)
5. JP3–13: Joint Doctrine for Information Operation, United States of America (2006)
6. FM100-06: Information Operations, Headquarters, Department of the Army, Washington, United States of America(1996)
7. Kurose, J.F., Ross, K.W.: Computer Networking, Addison Wesley, 4th edn. United States of America (2008)
8. Waltz, E.: Information Warfare: Principles and Operations. Artech House (1998)
9. FM3-13: Information Operations: Doctrine, Tactics, Techniques, and Procedures, Headquarters, Department of the Army, Washington, United States of America (2003)
10. Erbschloe, M.: Physical Security for IT. Elsevier Digital Press, United States of America (2005)
11. Tikk, E.: National Defense Policies for Cyber Space – Background and Effect of the Estonian Cyber Attacks. Academia Militar, Lisboa (2008)
12. Tikk, E., et al.: Cyber Attacks Against Georgia: Legal Lessons Identified, NATO Unclassified Report v1.0, Cooperative Cyber Defense Centre of Excellence, Tallin, Estonia (2008)
13. Krektel, B., Bakos, G., Barnett, C.: Capability of the People's Republic of China to Conduct Cyber Warfare and Computer Network Exploitation, Northrop Grumman Corporation, Report, United States of America (2009)
14. Alberts, D.S., Garstka, J.J., Stein, F.P.: Network Centric Warfare: Developing and Leveraging Information Superiorit, Washington, United States of America. CCRP Publication Series (1999)
15. Alberts, D.S., et al.: Understanding Information Age Warfare, Washington, United States of America. CCRP Publication Series (2001)

16. Hutchinson, W.: The Changing Nature of Information Security. In: 1st Information Security Management 2003, Australian (2003)
17. Chesla, A.: Information Security: A Defensive Battle. Information Security Journal: A Global Perspective 12(6), 24–32 (2004)
18. Laudon, K.C., Laudon, J.P.: Management Information Systems, 9th edn. Prentice Hall, United States of America (2006)
19. Pfleeger, C.P., Pfleeger, S.L.: Securiy in Computing, 9th edn. Prentice Hall, United States of America (2007)
20. Harris, S.: CISSP All-in-One Exam Guide, 4th edn. McGraw-Hill, New York (2008)
21. ISO/IEC13335-1: Information technology- Security techniques-Management of information and communications technology security. Part 1: Concepts and models for information and communication technology security management (2004)
22. Hong, K., et al.: An integrated system theory of information security management. Information Management and Computer Security 11, 243–248 (2003)
23. COBIT4.0: Control Objectives – Management Guidelines – Maturity Models, IT Governance Institute, United States of America (2005)
24. Vermeulen, C., Von Solms, R.: The information security management toolbox-taking the pain out of security management. Information Management and Computer Security 10(2/3), 119–125 (2002)
25. Finne, T.: A conceptual framework for information security management. Computers & Security 17(4), 303–307 (1998)
26. Nnolim, A., Steenkamp, A.: An Architectural and Process Model Approach to Information Security Management. Information Systems Education Journal 6, 31 (2008)
27. von Solms, B.: Information security—a multidimensional discipline. Computers & Security 20(6), 504–508 (2001)
28. Kajava, J., et al.: Information Security Standards and Global Business. Industrial Technology, 15–17 (2006)
29. Ma, Q., Johnston, A., Pearson, J.: Information security management objectives and practices: a parsimonious framework. Information Management & Computer Security 16(3), 251–270 (2008)
30. Baskerville, R.: Information systems security design methods: implications for information systems development. ACM Computing Surveys (CSUR) 25(4), 375–414 (1993)
31. Eloff, M., Von Solms, S.: Information security management: a hierarchical framework for various approaches. Computers & Security 19(3), 243–256 (2000)
32. Kritzinger, E., Smith, E.: Information security management: An information security retrieval and awareness model for industry. Computers & Security 27(5-6), 224–231 (2008)
33. Broderick, J.: ISMS, security standards and security regulations. Information Security Technical Report 11(1), 26–31 (2006)
34. Humphreys, E.: Information security management standards: Compliance, governance and risk management. Information Security Technical Report 13(4), 247–255 (2008)
35. ISO/IEC27002: Information Technology-Security Techniques-Code of Practice for Information Security Management (2007)
36. ISO/IEC13335-4: Information technology- Guidelines for the management of IT Security. Part 4: Selection of safeguards (2000)
37. ISO/IEC13335-5: Information technology-Guidelines for the management of IT Security. Part 5: Management guidance on network (2001)
38. Von Solms, R.: Information security management: why standards are important. Information Management and Computer Security 7, 50–57 (1999)
39. NIST-SP800-53: Information Security (2007)

40. NIST-SP800-42: Computer Security – Guideline on Network Security Testing (2001)
41. Barafort, B., Humbert, J., Poggi, S.: Information Security Management and ISO/IEC 15504: the link opportunity between Security and Quality (2006)
42. ISO/IEC15408: Information Technology-Security Techniques-Evaluation Criteria for IT Security (2005)
43. Alberts, C., Dorofe, A.: OCTAVE – Method Implementation Guide Version 2.0, Carnegie Mellon, Software Engineering Institute, United States of America (2001)
44. Huang, S., Lee, C., Kao, A.: Balancing performance measures for information security management. Industrial Management & Data Systems 106(2) (2006)
45. Martins, J.C.L., Santos, H.M.D.d., Nunes, P.V.: Security Framework for Information Systems. In: 8th European Conference on Information Warfare and Security, Lisboa (2009)
46. Farn, K.J., Lin, S.K., Fung, A.R.W.: A study on information security management system evaluation - assets, threat and vulnerability. Computer Standards & Interfaces 26(6), 501–513 (2004)

OTM Machine Acceptance: In the Arab Culture

Abdullah Rashed and Henrique Santos

R & D Algoritmi Centre, University of Minho, Guimarães, Portugal
{rashed,hsantos}@dsi.uminho.pt

Abstract. Basically, neglecting the human factor is one of the main reasons for system failures or for technology rejection, even when important technologies are considered. Biometrics mostly have the characteristics needed for effortless acceptance, such as easiness and usefulness, that are essential pillars of acceptance models such as TAM (technology acceptance model). However, it should be investigated. Many studies have been carried out to research the issues of technology acceptance in different cultures, especially the western culture. Arabic culture lacks these types of studies with few publications in this field. This paper introduces a new biometric interface for ATM machines. This interface depends on a promising biometrics which is odour. To discover the acceptance of this biometrics, we distributed a questionnaire via a web site and called for participation in the Arab Area and found that most respondents would accept to use odour.

Keywords: Technology acceptance, ATM, interface, Arab culture.

1 Introduction

In spite of biometrics adoption for the mainstream of authentication technologies, user acceptance is not yet resolved [5]. Moreover, one of the main reasons why the security systems fail is human misbehaviour [6]. Many people think that new technologies will be easily adopted [20]. However, some sensitive cases contracted with that assumption, as the first mechanical cash issuer was developed and installed, but removed after six months due to the lack of customer acceptance [14]. Therefore many studies have been carried out to discover the interface acceptance in different cultures especially the western culture. Arabic culture lacks to these kinds of studies with few publications in this field.

TAM and other models are used to measure the technology acceptance level. It depends on the perceive usefulness and ease of use to measure the intention to use the technology, as shown in figure 1.

Users need to access many technology tools. Passwords are the principal authentication technique, but its vulnerabilities are very important, especially for saving critical data or for massive utilization. In addition to that, user practices cannot be policed [19]. Users select easy guessed passwords [8] or something related to their daily life [4] and that is a trade-off between the security and usability and memorability. This problem can be solved using biometrics techniques [9].

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 131–138, 2010.

Fig. 1. Technology Acceptance model [20]

Moreover, biometrics is applied to effectively fight the identity theft [13]. Biometrics mostly have the characteristics needed for acceptance such as easiness and usefulness that formulate the essential pillars of TAM [15]. Therefore using biometric methods to user authentication is on increase [13]. Financial services are considered the key market for the biometrics industry [4]. Biometric implementation can be tested in the self-service environment, especially the ATM [4].

This paper introduces a new biometric interface for ATM machines, based on human odour. To study the acceptance of this biometrics in Arab countries, we distributed our questionnaire via a web site and called for participation.

The rest of the paper is organised as follows. In section 2; we overview the previous studies as literature review and address the problem statement. In section 3 we demonstrate our methodology and discussion. We conclude and present future work in section 4.

2 Literature Review

Al-Gahtani overviewed both TAM and the unified theory of acceptance and use of technology (UTAUT). They tried to validate them in Saudi Arabia. They hypothesized and tested the similarities and differences between the North American and Saudi validations of UTAUT in terms of cultural differences that affect the organisational acceptance of IT (information technology) in the two societies. They examined the relative power of a modified version of UTAUT in determining 'intention to use' and 'usage behaviour'. Their findings revealed that performance expectancy had a positive effect on intention, but they found no interacting effect with performance expectancy and either gender or age on intention. In addition, they found that effort expectancy did not have a significant effect on intention in the presence of interactions with the moderating variables. They found a negative interaction between effort expectancy and experience on intention. They claimed that increasing years of experience with computers, force ease of use to become less important in predicting Saudi's behavioural intentions. They thought that Saudis culture had high power distance with strong association subjective norm and behavioural intention [1].

Twati studied the cultural norms and beliefs within multi-national organisations in two regions: Arab countries on South Africa (Libya) and Arab countries on the Persian Gulf (Kuwait, Oman, Saudi Arabia, and United Arab Emirates). The study revealed that the two regions were not homogeneous. In addition, the study showed

that age, gender, and education levels are factors contributing to the success of MIS (Management Information Systems) adoption in the two regions. Furthermore, the study showed differences in organisational cultures that impact upon MIS adoption in both regions. The Persian Gulf region was dominated by an adhocracy culture that values the adoption of MIS, whereas the North Africa region was dominated by the hierarchy culture type that favours a centralised management style, which impacts negatively on MIS adoption. The Persian Gulf region did not show any significant effect of technology acceptance variables. However, in the North Africa region, technology acceptance played a vital role in MIS adoption [21].

Loch studied the culture-specific inducements and impediments to use the Internet in the Arab countries. They tried to measure how technology affects the acceptance of the Internet and how social norms affect the acceptance of the Internet. Their findings identified how culture can both inhibit and encourage technological innovation and how Arab culture could move their economies more quickly into the digital age [12].

Rose and Straub examined technology acceptance in five Arab cultures, three Asian (Jordan, Saudi Arabia, and Lebanon), two African countries (Egypt and the Sudan). They examined the ease of use and perceptions of usefulness. They studied the role of the two factors on actual usage and perceptions of usefulness to mediate the effect of perceptions of ease of use, on actual usage. Their findings were consistent with the majority of TAM findings in the US [18].

Conventry studied user acceptance and usability of biometric authentication in self-service places. Their research conducted biometric techniques, especially iris verification technology, at the Automated Teller Machine (ATM) user interface [4].

Dhamija [6] presented a Déjà vu focused on the human ability to recognise a previous seen image. They found it easier and more reliable than other approaches that depend on recalling passwords and PIN. They built a prototype for the system and conducted a study to compare it with password and PIN authentication. Their results showed that 90% of all participants succeeded in the authentication tests using Déjà vu while only about 70% succeeded using passwords and PINs. Their findings indicate that Déjà vu had potential applications, especially where text input might be hard (e.g., PDAs or ATMs), or where passwords will be infrequently used (e.g., web site passwords).

Rashed discussed the acceptance of odour as authentication technology among young people. They found that odour as authentication tool was accepted [15].

Dogs use their noses to recognize things via odour. When they search they can use their memory to remember as smell sense is linked to memory and emotion [3].

For the human beings; smell is not used. The reasons may be summarised in the lacking to the research and the IT tools to enable these devices to work [3]. Moreover this field is under development [10] so it is much less well understood then other biometric techniques [3].

Technology continues to evolve and improve, so more work is required to address the usability issues which will be the key to successful implementation biometrics within a general public application such as banking. Biometrics and smart cards continue to make the headlines in the government ID sector, where they are being applied for applications such as national ID, healthcare, and driving licenses [2]. The design and development of systems that include biometric devices pose specific challenges to integrators because of the interplay of technical and social factors [16]. Many biometrics have been applied around the world. Face recognition is used by a bank for authentication in internal communication [13].

Three patents related to odour machines are registered, as described in [15] who discussed the acceptance of odour in ATM machines as authentication technology among European young people. In spite of the security worries, odour would be accepted as authentication tool.

In addition; Wang introduced an odour interface invention to solve the problem of camera shortages. The invention provides the necessary interface for digital cameras, allowing users to add desired signature associated to odour to the selected digital image files. When browsing the digital image files containing odour information, the odour can be dispersed through an odour dispersing unit. Moreover, there is new a device (OdoReader) developed at the University of Bristol and Norman Ratcliffe from the University of the West of England. It can sniff out the presence of disease by smell [22]. It diagnoses Clostridium difficile, which may cause severe diarrhoea, especially amongst hospitalised patients [17].

A comparison between five beagles and five electronic devices was carried out to detect termites in wood. The results showed that beagles performed best for blocks containing 50 or more termites [11].

The American military delivered bomb-sniffing dogs to Iraq under pressure for using equipment that may be ineffective in finding explosives [7].

Moreover [10] introduced the electronically noses (ENoses) model as an odour recognition device.

Financial services are considered the key market for the biometrics industry [4]. To solve the problem of accessing the technology tools with easy way, we suggest using odour as authentication tool. We propose that odour can be used as an authentication method in banking and may improve the ATM machines concerning its security performance, Such ATM would be called odour ATM (OTM).

Odour automated teller machine (OTM) is supposed as a computerised telecommunications device that provides the customers of a financial institution with access to financial transactions in a public places without the need for a human clerk or bank teller. The customer is identified by odour.

The proposed model is an OTM machine (Odour ATM) exactly the same ATM in addition to sensors that detect the smell and the associated module for improved authentication.

We suppose that OTM would allow only one customer to use it (at time) at the same time and be alone with some special cabinet that ensures that.

The outline of the OTM process would require:

- A place for only one person.
- Odour sensors.
- Special keyboard (fingerprint scanner)

  Before starting developing further the biometric technique or the OTM technology, it is recommended to evaluate if this interface would be accepted by users, as stated before.

3 Methodology and Discussion

A questionnaire was distributed via website. We selected e-mail as the target technology. It has been studied in the existing literature [15].

The constructs used existing scales from previous studies. Questionnaire contained 9 statements and fully anchored 5-point Likert scales were used with end points being for one “extremely disagree” and five “extremely agree”.

Because the majority of Arab people are not proficient in English, the instrument was developed in English and was translated into Arabic. Chat and emails were used to clarify the ambiguity that respondents faced.

We received 62 responses, only two of them were in English and 60 responses were in Arabic. The main findings are:

Age of the respondents was within the interval [20-30] that represents youth people as shown in figure 2.

15-19	20-30	31-39	40-50	> 51
5	41	13	2	1

Fig. 2. Ages of the respondents

1. Education level: most of the respondents obtained high education as shown 3

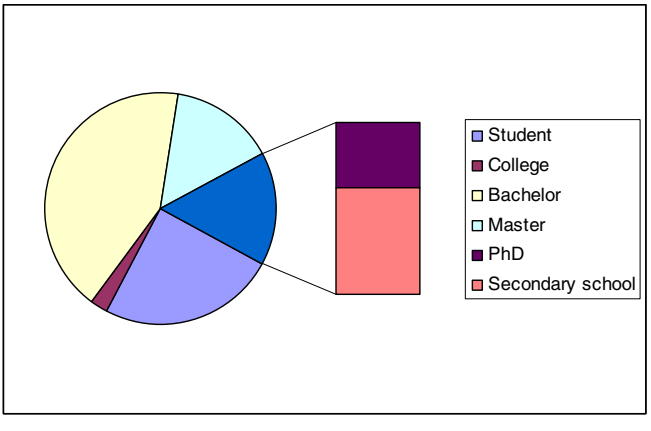


Fig. 3. Education level of the sample

2. The specialization of the sample: most of them are information technology users, as shown in figure 4.
3. Figure 5 shows that 31 of the participants (50%) found it easy to use odour as authentication system in ATM machine, whereas 8 of them (13%) thought it would not be easy.

- 37 respondents (60%) stated that odour as authentication system in ATM machine is a good idea.
- 37 respondents (60%) pointed out that the technique would improve their performance and 36 (58%) found it would enhance their effectiveness in life.
- 32 respondents (52%) intended to use odour as authentication system.

- All respondents (62, which represent 100%) have not ever used odour as authentication system.
- The majority of the respondents which represents 59% of the sample) would use it frequently if it would be available.

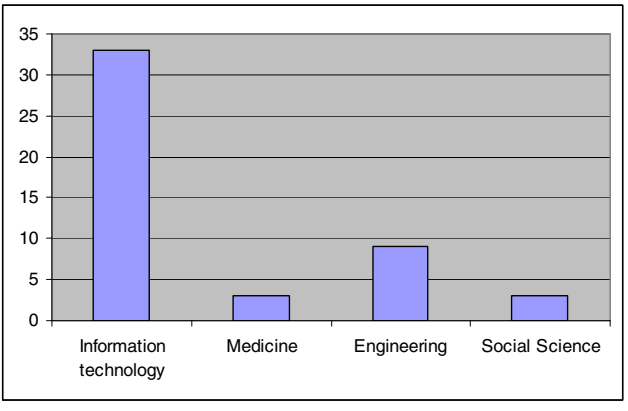


Fig. 4. Major of the sample

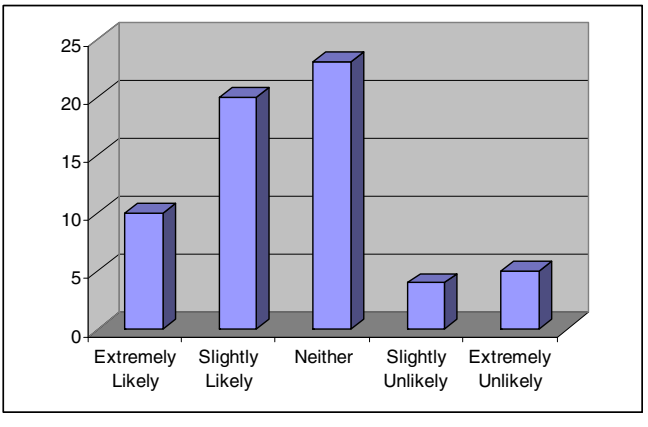


Fig. 5. Easiness of odour in ATM machine

4 Conclusion and Future Work

Most of the respondents were youth and they found it a good and useful idea to use odour as an interface in authentication system. Moreover, most of them have not used odour based technique before and they reported they would use it frequently, if it would be available. We have seen the international trend to use odour in agricultural, military and medical purposes. We think that the problem is how we could present the odour as interface. We think that odour can be used as an authentication method. It is a

challenge to apply this approach due to the lack of studies concerning the precision and scalability of this technique – these aspects require more research. Users think it is easy and needs to be strengthening with other approaches that enhance its performance.

Acknowledgement. Authors would like to thank http://www.estebyans.com web site for their cooperation.

References

1. Al-Gahtani, S., Geoffrey, S., Hubona, G., Wang, J.: Information Technology (IT) in Saudi Arabia: Culture and the Acceptance and Use of IT Source Information and Management 44(8), 681–691 (2007)
2. Beefing up security with biometrics. Card Technology Today, 14–15 (May 2008)
3. Brewster, S., McGookin, D., Miller, C.: Olfoto: Designing a Smell-based Interaction. In: Proceedings of the SIGCHI conference on Human Factors in computing systems, Montréal, Québec, pp. 653–662 (2006) ISBN:1-59593-372-7
4. Conventry, L., Angeli, A., Jonson, G.: Usability and Biometric Verification at the ATM Interface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Archive, Ft. Lauderdale, Florida, USA, pp. 153–160 (2003)
5. Coventry, L.: Biometrics, self-service and the user. Biometric Technology Today, 7–9 (November/December 2004)
6. Dhamija, R., Perrig, A.: Déjà Vu: A User Study Using Images for Authentication. In: 9th USENIX Security Symposium, Denver, Colorado, US, p. 4 (2000)
7. Fox News: U.S. Rushing Bomb-Sniffing Dogs to Iraq amid Skepticism over Equipment (2010), `http://www.foxnews.com/story/0,2933,584801,00.html`
8. Gong, L., Lomas, M., Needham, R., Saltzer, J.H.: Protecting Poorly Chosen Secrets from Guessing Attacks. IEEE J. on Selected Areas in Communications 11, 648–656 (1993)
9. Heckle, R., Patrick, A., Ozok, A.: Perception and Acceptance of Fingerprint Biometric Technology. In: Proceedings of the 3rd symposium on Usable privacy and security, Pittsburgh, Pennsylvania, USA. ACM International Conference Proceeding Series, vol. 229, pp. 153–154 (1996) ISBN:978-1-59593-801-5
10. Korotkaya, Z.: Biometric Person Authentication: Odor (2009), `http://www.it.lut.fi/kurssit/03-04/010970000/seminars/Korotkaya.pdf`
11. Lewis, V.R., Fouche, C.F., Lemaster, R.L.: Evaluation of Dog-Assisted Searches and Electronic Odor Devices for Detecting the Western Subterranean Termite. Forest-products-journal (USA) 47(10), 79–84 (1997)
12. Loch, K., Straub, D., Kamel, S.: Diffusing the Internet in the Arab world: the role of social norms and technological culturation. IEEE Transactions on Engineering Management 50, 45–63 (2003)
13. McIntosh, D.: Biometrics – a fad or the future? Biometric Technology Today, 9–11 (2009)
14. MIT School of Engineering: Inventor of the Week: Range Estimation Trainer (2003), `http://web.mit.edu/invent/iow/simjian.html`
15. Rashed, A., Santos, H.: Odour User Interface for Authentication: Possibility and Acceptance: Case Study. In: The International MultiConference of Engineers and Computer Scientists 2010 (IMECS 2010), The 2010 IAENG International Conference on Bioinformatics, Hong Kong (2010)

16. Rejman-Greene, M.: A Framework for the Development of Biometric Systems. Biometric Technology Today, 6–8 (2003)
17. Rooney, S.: University of Bristol OdoReader Could Save Health Services Millions by Sniffing Out Stomach Bugs (February 2010), http://www.medicalnewstoday.com/printerfriendlynews.php?newsid=178283
18. Rose, G., Straub, D.: Predicting General IT Use: Applying TAM to the Arab World. Journal of Global Information Management 6(3), 39–46 (1998)
19. Skaff, G.: An alternative to passwords? Biometric Technology Today, 10–11 (2007)
20. Tibenderana, P., Ogao, P.: Acceptance And Use of Electronic Library Services in Ugandan Universities. In: International Conference on Digital Libraries, Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries, Pittsburgh, PA, USA (2008)
21. Twati, J.M.: Societal and Organisational Culture and the Adoption of Management Information Systems in Arab Countries, Ph.D. theses, Griffith University (2009)
22. Wang F.: Patent application title: Device and Method for Setting Odor, Inventors: Agents: Rosenberg, Klein & Lee Assignees: Altek Corporation Origin: Ellicott City, MD US IPC8 Class: AG06F1730FI USPC Class: 7071041 (2009), http://www.faqs.org/patents/app/20090132596#ixzz0eTfz5AQY

A Study on the Interrelations between the Security-Related Antecedents of Customers' Online Trust

Hamid Reza Peikari*

Graduate School of Business, University Kabangsaan Malaysia
(National University of Malaysia- UKM), 43600 Bangi, Malaysia
omid726@yahoo.com

Abstract. Despite the wide attention of previous studies to explore the influence of different security-related factors on customers' online trust, the interrelations between such factors and their direct and indirect influences on customers' trust have been neglected. This study investigates the direct and indirect interrelations between the factors authentication, encryption, technical protection and externally provided assurances including third party security and privacy seals with customers' trust in the business-to-customer (B2C) environment. The data was collected from 238 respondents and after the test of reliability and validity of the scale, the hypotheses were tested using structural equation modeling. The results showed that customers' perception of encryption and authentication mechanisms implemented by a Website have a positive significant influence on their perceived technical protection while technical protection was found to significantly influence customers' trust to the Website. However, the analysis did not find any relation between the third party assurance and customers' trust, indicating that despite the high expenses companies involve to obtain such assurances from reputed third parties, such mechanisms and assurances do not have any direct or indirect significant influence on customers' trust; which raises questions on the value of such mechanisms .finally, after discussing the findings and implication of this study for both academic and business worlds, suggestions for future studies were made to have a better understanding of the dimensions of the interrelations between the security-related factors.

Keywords: Trust, technical protection, Authentication, Encryption, Externally provided assurance, Security-related factors.

1 Introduction

The development of Information and Communications Technology (ICT) and particularly Internet has created a revolution in the life styles and standards of living of human community. Adding the letter "e" to any of today's practices and entities has become a norm to describe an online process through Internet such as e-citizen, e-city, e-service,

* Corresponding author.

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 139–148, 2010.

e-purchase and e-government. With the rapid pace of Internet-based applications development and their various advantages, governments, businesses and individuals seek to use e-business technologies and applications in their day to day activities. However there are many issues and challenges related to the global diffusion of e-commerce to be addressed by both academic and business worlds. One of these issues is customers' trust [10, 26, 32]. Jarvenpaa et al. [17] believe that while in the early stages of Internet application, a user's concerns about trust were more about the performance of technologies, in the late stages of online technologies; they are more concerned on how firms have implemented their technologies. Among them, still the circumstances under which, customers are willing to involve a transaction with a Website is still unknown to researchers [5]; for instance, which factors can influence customers to trust a website and provide their personal and credit card information to make a purchase.

Security and privacy features of a Website have been cited as one of the determinant factors of customers' trust [6, 8, 15, 25]. Given the fact that the studies investigating the security perceptions of customers are limited in their understanding of the internet security measures, it is important to establish the measures of internet security perception and its relations to perceived trust in online transactions [9]. Further more, the interrelations between different dimensions of security solutions of Website have not been fully investigated and understood by researchers and only the influence of such factors on customers' trust have been explored [21, 22, 28].

This research intends to investigate the interrelations between some of the security related factors and also their direct and indirect influence on the customers' trust in a B2C Website. The findings of this study will contribute in the academic world by improving the current models and understanding the interrelations exist between such factors and trust. The findings also help companies and their managers to have a better understanding and consideration of the security related antecedents of trust, which enables them to formulate and implement more efficient and effective e-commerce strategies and design.

2 Consumer Trust in E-Commerce

Barmall et al. [4] believe that despite the wide attention of social and business communities and entities toward trust, yet, there is no consensus on its definition. Yousafzai et al. [33] defined trust as "a function of the degree of risk involved in the e-banking transaction, and the outcome of trust is proposed to be reduced perceived risk, leading to positive intentions towards adoption of e-banking" (p 847). Trust production mechanisms have been categorized into different facets such as process-based, characteristic-based, and institutional-based trust [21, 34], dispositional or personality-based trust, effect-based or cognition-based trust, and institutional-based trust [3], relational and technological trust [23], initial trust, knowledge-based trust, calculative based trust, personality based, institution based and cognition based trust [24].

According to the theory of planed behavior [2] and the theory of reasoned action [1, 13] human behavior is influenced by his behavioral intention and behavioral intention is a function of attitude. According to Njite and Parsa [27], customers' online trust depends on their perception that whether online merchants perform particular activities that they are supposed to or not such as merchant's commitment to maintain

the confidentiality of the data and information they receive from or send to their customers. Therefore, it can be argued that customers' attitude toward the merchants' security measures and solutions plays an important role on customers' perceived trust to the merchant.

The factors related to confidentiality and security of customers' data and information are one of the main facets of customers' trust to a Website. Such factors aim to deal with security threats associated with online transactions. Yousafzai et al. [33] refer to security threat as "a threat which creates circumstance, condition, or event with the potential to cause economic hardship to data or network resources in the form of destruction, disclosure, modification of data, denial of service, and/or fraud, waste, and abuse" (p 853) such as unauthorized access to confidential data and information, network or data transaction attacks or similar fraudulent activities.

Previous studies have referred to different confidentiality-related factors from different perspectives. For instance, some authors have used the broad terms "security" and "privacy" [9, 12, 14, 28] while some others have been more detailed and specific on this regard and have referred to such factors as different technological solutions by specifying the nature and application of such factors [5, 18, 30], or externally and internally provided assurances and policy statements [3, 21, 22].

Except Kim et al. [18] who studied the interrelations between the two independent factors "technical protection" and "security statement" and the mediator factor perceived overall security with customers' trust, other researchers have not attempted to study the interrelations might exist between the security-related factors and only the relations between such factors with trust have been investigated. The research executed by Kim et al. [18] however, was very limited in its scope on this regard, as it did not considered all the security-related factors and only a few of them were explored. Therefore, there is a need to study the interrelations between such factors to have a better understanding of their influence on each other and their direct and indirect influences on customers' trust.

This study refers to such factors as the detailed specific technological solutions to study the relations between the specific factors and their influence on trust. Four different factors identified as the specific security-related solutions in this study are authentication, encryption, technical protection and externally provided assurance.

3 Technical Protection

Literature in computer sciences classifies information security issues from different perspectives. For example, based on the point of origin, it can be classified as those occurred at the point of origin, during data transition, at the destination or after storage or can be classified based on the goal of the attack such as loss of confidentiality, availability or integrity [9].

Kim et al. [18] define Website technical protection as the technical mechanisms employed to protect consumers' transaction security and argue that a reliable level of privacy, integrity and stability can improve the customers' level of perceived trust and security in the context of e-payment systems. This study refers to this factor as the technological capabilities of a Website to ensure the confidentiality of the exchanged data and information. Technical protections of Websites have been categorized into

objective and subjective type [19]. Since customers find it difficult to evaluate the protection mechanisms and technologies from an objective perspective, they use their perceptions and expectations toward the functionality of such solutions.

This study argues that there are differences between the customers' perceived technical protection measures implemented by a Website and the assurances and evidences demonstrating the Website commitment and capability to protect customers' sensitive data from unauthorized access. Therefore, this study improves the model proposed by Belanger et al. [5] and refers to antecedents of trust from two main dimensions: the first dimension refers to the technologies implemented by the website such as encryption and authentication solutions while the second dimension will be the externally provided assurances of the Website commitment and competence to maintain the confidentiality of the data. Given the above discussion, it is suggested that:

H1: customers' overall technical protection mechanisms of a Website have a positive significant influence on their trust to the Website.

4 Encryption

Encryption, according to Chellappa and Pavlou [9] is using complex mathematical algorithms and keys to encode information and data from their original form (plaintext) to an incomprehensible form (cipher text). Only authorized parties have the key to decrypt the information to the proper and understandable format. Encryption ensures that the information provided to the Website will not be interrupted or altered by any party [18] even if they can access it.

It was found that encryption technologies influence customers' perceptions toward security and trust in e-payment systems [18]. Since customers evaluate the security mechanisms of a Website from an objective perspective using their personal perceptions and not based on the subjective technological angle [9], therefore, what matters is the customers' perception toward the encryption mechanism of a Website. Customers' perception of the functionality and effectiveness of encryption mechanisms of a Website influences their perception toward the functionality of the technical protection mechanisms of the Website. In other words, the more efficient the encryption solutions of a Website works, the more efficient would be the overall technical protections of the Website. Therefore, it can be hypothesized that:

H2: Customers' perceived encryption mechanism deployed by a Website is positively related to their perceived technical protection measures of the Website.

5 Authentication

Authentication techniques are the most accepted and widely used security solutions in IT [15]. In the context of e-commerce, it refers to the mechanisms through which a trusted third party guarantees that the online merchant is the one that claims to be [9]. In fact it is a mechanism to determine whether someone is who s/he has declared to be.

Authentication technologies as crucial solutions in enhancing customers' confidence in online transactions, guarantee the confidentiality of the information by preventing unauthorized parties to capture or infer it [18]. This study measures au-

thentication mechanisms of a website from the customers' perspective. Authentication solutions of a Website, ensures that the data and information sent by the customers will receive only to the right party (merchant). Therefore, such techniques enhance customer perception toward the efficiency of overall technical protection of the Website. Given this fact, it can be hypothesized that:

H3: Customers' perceived technical protection of a Website is positively influenced by their perceived authentication mechanism employed by the Website.

While authentication mechanisms ensure that the data receives the right party, encryption techniques ensure that only the right party can understand the data received; so even in case the data is accessed by an unauthorized party, it can not be understood due to the decrypted format of the message. Therefore, though the functionalities of encryption and authentication technologies are different, their ultimate objective and result are similar. Thus, it can be suggested that:

H4: There is a positive relation between customers' perceptions toward encryption and authentication mechanisms of a Website.

6 Externally Provided Assurance

This study suggests that other than the perception of customers' toward the technical protection measures implemented by a website, there is a second set of factors called as the externally provided assurances, which refers to third party evaluations of the efficiency of protection mechanisms of a Website and the competence and commitment of the Website to maintain the confidentiality of data. In other words, while the factor perceived technical protection refers to customers' perception and expectations toward the technologies implemented by a Website to guarantee the data confidentiality, the second factor- externally provided assurance- refers to the assurances provided by external parties ensuring the commitment and competence of the Website on this regards.

Some authors have classified such assurances into two groups, internally provided by the Website and externally (third party) provided assurances by an independent reputable body after testing and evaluating the Website practices and technology [3]. They speculate that such assurances are important for especially a new established Website as customers do not have any experiences with a new Website and find it difficult to trust it; but presence of such assurances on such websites may establish trust in customers. Their findings indicate that externally provided assurances are no as important as internally provided assurances in establishing and enhancing customers' perceived trust to a Website. However, they did not study the indirect influence of these factors on trust, which might have a significant indirect influence on that variable.

Doney and Cannon [11] assert that online shoppers tend to share private and/or sensitive data and information with other online entities such as online merchants through a third-party mechanism. Many Websites use third party assurance seals, ensured by a reputable third party to show their capability, competence and commitment toward maintaining security and privacy of their users. Bahmanziari et al. [3] speculate that such assurances are provided by third parties only after an independent comprehensive assessment of the Website and its related activities. When the competence

and commitment of the Website is ensured, a certificate is issued by the third party as the evidence of such capability and commitment. The third party security and privacy seal indicate that the merchant operating through the Website has adequate policies, technologies and capabilities to ensure the privacy and security of its users.

Previous studies have referred to such factors as the antecedent variables of online trust from a specific perspective [5, 22, 30] or more broad classifications [3, 21]. Moreover, as discussed above, such assurances are given to a Website after a comprehensive evaluation of the overall technical protection mechanisms of the Website. Therefore, it is expected that not only influences customers' trust to the Website, but also influences their perceptions toward the technical protection of the Website. Therefore, it is suggested that:

H5: Externally provided assurances of a Website positively influences customers' trust in a Website.

H6: Externally provided assurances of a Website positively influences customers' perception of technical protections of the Website.

7 Research Method

This research employed a self-administered questionnaire with structured closed ended questions as the data collection method and the data was collected in Malaysia. As shown in the Table 1, some of the scale items were adopted and adapted from previous published studies while some other items were developed by the researcher and the data collected from of 238 respondents was used for final data analysis. The sample size is consistent with the sample size used in some studies in this context [16, 20, 32]. Respondents were asked to fill out the questionnaire with a 'familiar e-commerce website' in mind and no specific Website was mentioned as their reference to eliminate the influence of the Website brand from their answers. This method ensures that respondents have sufficient knowledge and familiarity with the Website of which they evaluate its features [7]. Moreover, asking respondents to answer the questions with any familiar Website in mind without mentioning any Website name is valid and acceptable as there are studies available by some researchers employed similar method [3, 7].

It was found that majority of the respondents (62.2%) are females. Moreover, it was found that 59.1% had the experience of more than one time online purchase. The self reported IT skills of the respondents indicate that 94.6% of them evaluate their IT skills as an average or above average level. The results of online purchase experience and IT skills of the respondents indicate that they have a relatively enough knowledge to understand and respond the questions.

To ensure the reliability of the scale, the test of Cronbach's alpha was performed by SPSS 16 and as shown in the Table 2, the scale was found highly reliable.

The validity of the scale was administered by the maximum likelihood estimation technique of confirmatory factor analysis (CFA). As shown in the Table 2, all the fitness values of the measurement model are above the cut-off level, suggesting a good model fit.

Table 1. The reliability of the final questionnaire

Variable	Source	No. of Items	Cronbach's Alpha
Technical Protection	[18], New	3	0.828
Ext. Assurance	New	3	0.934
Authentication	[9], New	3	0.846
Encryption	New	2	0.826
Trust	[18]	4	0.888

Table 2. Fit Indices for measurement and structural model

Fit Indices	Recommended value [31]	Result value (Measurement Model)	Result value (Structural Model)
$\chi 2$	NA	117.089	130.039
d.f.	NA	80	84
$\chi 2$/d.f.	≤ 3	1.464	1.548
GFI	≥ 0.9	0.940	0.933
NFI	≥ 0.9	0.950	0.945
NNFI	≥ 0.9	0.078	0.974
CFI	≥ 0.9	0.984	0.980
RMSEA	≤ 0.05	0.044	0.048
RMSR	≤ 0.05	0.035	0.047

Convergent validity was tested following the guidelines proposed by Chang and Chen [8] and all the items loading with their corresponding latent variables were found significant (t-values > 4, p-values< 0.001) while all the items loadings were above 0.70. Discriminant validity was tested following the guidelines suggested by Salisbury et al. [29] and all the paired correlations between the latent variables were found less than 1, indicating the discriminant validity of the scale. Therefore, the validity of the questionnaire was ensured.

8 Analysis and Findings

In order to test the hypothesis, structural equation modeling (SEM) was applied using AMOS 5. As shown in Table 2, the goodness of fit was tested for the structural model and the results were found above the recommended level, indicating the goodness of the model fit.

To test the hypotheses, path analysis technique was applied and it was found that the influences of encryption on technical protection ($P<0.05$, $t =2.22$), authentication on technical protection ($P<0.01$, $t =3.06$) and technical protection on customers' trust ($P<0.001$, $t =6.37$) are significant and positive. Therefore, the hypotheses H1, H2 and H3 are supported. However, the hypotheses H5 and H6 related to the influence of externally provided assurance on customers' perceived trust and technical protection had to be rejected since the relations were not significant. Furthermore, the study found that there is a positive correlation between the variables encryption and authentication ($P<0.001$, $t =8.09$). Moreover, the squared multiple correlation value of the

variable technical protection indicates that more than 89% the variance in this variable is explained by the above antecedent factors. In other words, the above security-related variables strongly predict customers' perceived overall technical protection mechanism of a Website.

9 Conclusion

The results confirm the findings of previous studies on the positive significant influence of overall technical protection on customers' trust in a new setting [5, 18, 28]. Furthermore, the findings confirm the results of previous studies which did not find any significant relation between externally provided assurances and customers' perceived trust to the Website [3]. In other words, there are no evidences that externally provided assurances have any significant direct or indirect influences on customers' trust. This point implies that despite the huge amounts companies invest to obtain externally provided assurances, there are serious questions and doubts on the effectiveness of such assurances to build customers' trust. This might be due to the reason that individuals with a non-technical knowledge of e-commerce and especially security related issues are not familiar with such assurances and do not establish their security-related trust to a Website based on such evidences.

Moreover, the findings contribute in the body of knowledge by proposing a new model on the interrelations between the determinant factors of overall technical protection which was not explored in previous studies. In other words, while the factors related to the technical protection of customers' data is one of the factors building customers' trust to the Website, customers' perceptions of the technological capabilities of the Website- such as authentication and encryption- can enhance customers' overall perceived technical protection mechanisms of the Website. The results imply that in order to enhance customers' perceptions of a Website overall technical protection, Websites and online merchants should efficiently and effectively communicate the Websites encryption and authentication technologies to their customers and target markets.

However, this research is not free of limitations. First limitation is that this study does not investigate the influence of internally provided assurances-such as protection policy statements- on the model variables and dimensions. Furthermore, the effects of some other Website features such as user interface elements on the model variables were not studied. Another limitation is that the role of customers' characteristics such as risk and trust propensity on the relations between the variables was not explored. Moreover, the data was collected in a developing country (Malaysia) and the participants' responses may have been influenced by factors such as their national culture, e-readiness or economic situation and a study on different countries especially in developed countries with higher e-readiness indices and better diffusion of e-commerce may show different results, especially on the variable "externally provided assurance". Therefore, it is suggested that future studies make a more comprehensive model to study the influence of variables interface design, internally provided assurances and customers' characteristics on the antecedent variables of overall technical protection. Moreover, a comparison study between developing countries and developed countries is suggested to comparison the results and probable differences.

References

1. Ajzen, I., Fishbein, M.: Understanding Attitudes and Predicting Social Behavior. Prentice-Hall, Englewood Cliffs (1980)
2. Ajzen, I.: From Intentions to Actions: A Theory of Planned Behavior. In: Kuhl, J., Becaman, J. (eds.) Action-control: From Cognition to Behavior, pp. 11–39. Springer, Heidelberg (1985)
3. Bahmanziari, T., Odom, M.D., Ugrin, J.C.: An Experimental Evaluation of the Effects of Internal and External E-assurance on Initial Trust Formation in B2C E-commerce. International Journal of Accounting Information Systems 10, 152–170 (2009)
4. Barmall, C., Schoefer, K., McKechnie, S.: The Determinants and Consequences of Consumer Trust in E-retailing: A Conceptual Framework. Irish Marketing Review 17(1), 13–22 (2004)
5. Belanger, F., Hiller, J.S., Smith, W.J.: Trustworthiness in Electronic Commerce: the Role of Privacy, Security, and site Attributes. Journal of Strategic Information Systems 11, 245–270 (2002)
6. Berthon, P., Pitt, L., Cyr, D., Campbell, C.: E-readiness and Trust: Macro and Micro Dualities for E-commerce in a Global Environment. International Marketing Review 25(6), 700–714 (2008)
7. Chang, H.H., Chen, S.W.: The Impact of Online Store Environment Cues on Purchase Intention; Trust and Perceived Risk as a Mediator. Online Information Review 32(6), 818–841 (2008)
8. Chang, H.H., Chen, S.W.: Consumer Perception of Interface Quality, Security, and Loyalty in Electronic Commerce. Information & Management 46, 411–417 (2009)
9. Chellappa, R.K., Pavlou, P.: Perceived Information Security, Financial Liability and Consumer Trust in Electronic Commerce Transactions. Logistics Information Management 15(5), 358–368 (2002)
10. Chen, Y., Barnes, S.: Initial Trust and Online Buyer Behaviour. Industrial Management & Data Systems 107(1), 21–36 (2007)
11. Doney, P.M., Cannon, J.P.: An Examination of the Nature of Trust in Buyer-Seller Relationships. Journal of Marketing 61, 35–51 (1997)
12. Eastlick, M.A., Lotz, S.L., Warrington, P.: Understanding Online B-to-C Relationships: An Integrated Model of Privacy Concerns, Trust, and Commitment. Journal of Business Research 59, 877–886 (2006)
13. Fishbein, M., Ajzen, I.: Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Addison-Wesley, Reading (1975)
14. Flavian, C., Guinaliu, M.: Consumer Trust, Perceived Security and Privacy Policy; Three Basic Elements of Loyalty to a Web Site. Industrial Management & Data Systems 106(5), 601–620 (2006)
15. Furnell, S.M., Dowland, P.S., Illingworth, H.M., Reynolds, P.L.: Authentication and Supervision: A Survey of User Attitudes. Computers & Security 19(6), 529–539 (2000)
16. Gefen, D.: E-commerce: the Role of Familiarity and Trust. The International Journal of Management Science 28, 725–737 (2008)
17. Jarvenpaa, S.L., Tractinsky, N., Vitale, M.: Consumer Trust in an Internet Store. Information Technology and Management 1(12), 45–71 (2000)
18. Kim, C., Tao, W., Shin, N., Kim, K.: An Empirical Study of Customers' Perceptions of Security and Trust in E-payment Systems. Electronic Commerce Research and Applications 9, 84–95 (2010)

19. Linck, K., Pousttchi, K., Wiedemann, D.G.: Security Issues in Mobile Payment from the Customer Viewpoint. In: Proceedings of the 14th European Conference on Information Systems (ECIS 2006), Goteborg, Schweden, June 12-14, pp. 1–11 (2006)
20. Liu, C., Marchewkaa, J.T., Lub, J., Yub, C.: Beyond Concern: A Privacy–Trust–Behavioral Intention Model of Electronic Commerce. Information and Management 42, 127–142 (2004)
21. Luo, X.: Trust Production and Privacy Concerns on the Internet; A framework Based on Relationship Marketing and Social Exchange Theory. Industrial Marketing Management 31, 111–118 (2002)
22. Mahmood, O.: Trust: From Sociology to Electronic Environment. Journal of Information Technology Impact 6(3), 119–128 (2006)
23. McCord, M., Ratnasingam, P.: The Impact of Trust on the Technology Acceptance Model in Business to Consumer E-commerce. In: International Conference of the Information Resources Management Association: Innovations Through Information Technology, New Orleans, USA, May 23-26 (2004)
24. McKnight, D.M., Cummings, L.L., Chervany, N.L.: Initial Trust Formation in New Organizational Relationships. Academy of Management. The Academy of Management Review 23(3), 473–490 (1998)
25. Miyazaki, J., Fernandez, K.: The Antecedents and Consequences of Trust in Online Purchase Decisions. Journal of Interactive Marketing 16(2), 47–63 (2000)
26. Mutz, D.: Social Trust and E-commerce. Public Opinion Quarterly 69(3), 393–416 (2005)
27. Njite, D., Parsa, H.: Structural Equation Modeling of Factors that Influence Consumer Internet Purchase Intentions of Services. Journal of Services Research 5(1), 43–58 (2005)
28. Roca, J.C., Garcia, J.J., Vega, J.J.: The Importance of Perceived Trust, Security and Privacy in Online Trading Systems. Information Management & Computer Security 17(2), 96–113 (2009)
29. Salisbury, W.D., Pearson, R.A., Pearson, A.W., Miller, D.W.: Perceived Security and World Wide Web Purchase Intention. Industrial Management & Data Systems 101(4), 165–176 (2001)
30. Suh, B., Han, I.: The Impact of Customer Trust and Perception of Security Control on the Acceptance of Electronic Commerce. International Journal of Electronic Commerce 7(3), 135–161 (2003)
31. Tung, F., Chang, S., Chou, C.: An Extension of Trust and TAM Model with IDT in the Adoption of the Electronic Logistics Information System in HIS in the Medical Industry. International Journal of Medical Informatics 77, 324–335 (2008)
32. Wang, C., Chen, C., Jiang, J.: The Impact of Knowledge and Trust on E-consumers' Online Shopping Activities: An Empirical Study. Journal of Computers 4(1), 11–18 (2009)
33. Yousafzai, S.Y., Pallister, G.H., Foxall, G.R.: A Proposed Model of E-trust for Electronic Banking. Technovation 23, 847–860 (2003)
34. Zucker, L.: Production of Trust: Institutional Sources of Economic Structure: 1840– 1920. In: Staw, B., Cummings, L. (eds.) Research in organizational behavior, vol. 8, pp. 53–111. JAI Press, Greenwich (1986)

Does Nationality Matter in the B2C Environment? Results from a Two Nation Study

Hamid Reza Peikari

Graduate School of Business, University Kebangsaan Malaysia
(National University of Malaysia- UKM), 43600 Bangi, Malaysia
omid726@yahoo.com

Abstract. Different studies have explored the relations between different dimensions of e-commerce transactions and lots of models and findings have been proposed to the academic and business worlds. However, there is a doubt on the applications and generalization of such models and findings in different countries and nations. In other words, this study argues that the relations among the variables of a model ay differ in different countries, which raises questions on the findings of researchers collecting data in one country to test their hypotheses. This study intends to examine if different nations have different perceptions toward the elements of Website interface, security and purchase intention on Internet. Moreover, a simple model was developed to investigate whether the independent variables of the model are equally important in different nations and significantly influence the dependent variable in such nations or not. Since majority of the studies in the context of e-commerce were either focused on the developed countries which have a high e-readiness indices and overall ranks, two developing countries with different e-readiness indices and ranks were selected for the data collection. The results showed that the samples had different significant perceptions of security and some of the Website interface factors. Moreover, it was found that the significance of relations among the independent variables ad the dependent variable are different between the samples, which questions the findings of the researchers testing their model and hypotheses only based on the data collected in one country.

Keywords: nationality, Website features, security, purchase intention.

1 Introduction

To make customers purchase from a Website, it should meet three different needs of customers: need for particular information, need for particular service and need for particular Website features [37]. Similarly, some studies found that Website features positively influence customers' purchase intention from a Website [17, 18, 19, 31, 42]. It was found that customers' needs of Website features are influenced by different factors such as customers' demography [28], customers' personality [36, 40] and type of service offered by a Website [38]. However, it is still unknown that whether individuals living in different countries have different Website feature needs or not. In other words, still it is unknown for researchers that whether the same Website features influence customers'

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 149–159, 2010.

purchase intention from a Website, regardless of the nationality of the customers or customers living in different countries and areas have different Website feature needs and perceptions. Therefore, this study intends to meet the following objectives:

1. To test if Website features influencing the purchase intention of individuals living in different countries differ from each other.
2. To empirically examine if the individuals living in different countries have different perceptions toward some of the Website features.

The contribution of this study is not related to its research model, but is related to the changes in the significance of the influence of the independent variables on the dependent variable based on the samples' nationality. From the academic perspective, this study will contribute in the body of knowledge by proposing a new factor –customers' nationality- which may contribute in customers' needs and perceptions of Website features leading into their online purchase intention. Furthermore, the findings improve our understanding of the global diffusion of e-commerce in the developing countries. The results of the study will contribute in the business world by introducing a new factor which should be considered in the formulation of their e-commerce strategies and design of their Websites.

2 Nationality and E-Commerce

It is believed that different nations have different cultures [8, 14, 15], national corruption [8] and e-readiness [20] which may influence citizens online behaviors. For instance, it was found that the perception of different nations in different courtiers toward issues related to online environment such as trust, security and online purchase are different [14, 15]. However, they just established their analysis and conclusion based on the mean variations among their samples and did not applied other analysis to find out if such a difference are real differences or just a result of the differences naturally occur among the mean values of any different samples. Moreover, they did not study if the determinant variables of online purchase intention differ among the nations. Furthermore, they collected data from four developed countries with closed e-readiness ranks while it is a need to study the issues related to the diffusion of e-commerce in the developing countries, where a few studies are available on this context [1].

3 Purchase Intention

According to Poddar et al. [34] "purchase intent refers to the likelihood that a user makes a purchase from a site" (p 444). Similarly, Alcaniz et al. [5] refer to purchase intention as "a mental state that reflects the consumer's decision to acquire a product or service in the immediate future" (p 649). According to the theory of planned behavior [3], and the theory of reasoned action (2, 23) human behavior is a reflection of his behavioral intention and behavioral intention is the result of his attitude. According therefore, when a customer intends to purchase from a Website, his/her intentions are the function of his/her attitudes toward the Website features [24].

According to technology acceptance model (TAM), Website features such as ease of use and perceived usefulness influence user's intention to purchase from the Website [17]. Kim et al. [27] believe that customer's intention to use e-commerce technologies and systems is influenced by both perceived security while trustworthiness and security are associated with each other. Wen [41] found that a well designed Website enhances customers' purchase intention. Therefore, factors such as Website interface elements and security features of a Website are important Website features positively influence customers' purchase intention from a Website.

4 Website Interface Features

Website design quality is the "users' evaluations of whether a web site's features meet users' needs and reflect the overall excellence of the web site" [10: 821]. Despite the wide attention of researchers to Website interface features, there is no consensus among them on the factors should be studied to measure the user interface features of a Website [41]. Different studies have referred to different factors as the Website interface features such as appearance, and consistent image [30], color appeal [16], ease of use, usefulness, enjoyment [25], graphics, text, navigation, information presentation, ease of use, search engine, usability [11], information organization and layout, information content, information design [29], content, navigation, design and structure, appearance and multimedia [32]. This study refers to three interface factors ease of use, information presentation and navigability features of a Website interface.

4.1 Ease of Use

Ease of use features of a Website have been defined as "the extent to which a customer feels that a website is easy to navigate" [9, p 4] and has been cited as one of the most important dimensions of interface design in TAM [17]. This Website feature has been also studied in the works of some authors [6, 13, 25, 26, 35, 43] which demonstrates its importance. It is argued that about two-thirds of online transactions are not completed due to lack of this factor in the Website design [9]. Therefore, there is a relation between ease of use features of a Website interface and customers' purchase intention from the Website. Furthermore, according to TAM, ease of use of a Website influences the users' acceptance and intention to use of the Website. Therefore, it is suggested that:

H1: customers' perceived ease of a use of a Website significantly influences their intentions to purchase from the Website.

4.2 Information Presentation

Unlike traditional markets, customers' point of experience with a merchant/company is not through the sales staff of the company, but through their Website. While in a physical store, customers can touch and evaluate a product and have a real-situation interaction with the sales people, in the online environment, customers can not have a real-situation interaction to receive the information they need. Therefore, one of the challenges for online customers is gaining their required information from a Website,

which highlights the importance of the organization and presentation of information in the Website.

Information presentation refers to the relevance of information presented in a Website with its purpose while having adequate usefulness, scope and depth [38]. This dimension has been cited as one of the important features of a Website interface in previous studies [11, 21, 29, 32, 38] and some authors have argued that customers decisions to acquire a product or service is positively influenced by the information presentation features of a Website. Therefore, it is suggested that:

H2: Information presentation features of a Website significantly influence customers' purchase intention from the Website.

4.3 Navigation

Navigation features of a Website have been cited in previous research as one of the important dimensions of Website interface elements [11, 21, 32, 38] and depend on characteristics such as the number and effectiveness of hyperlinks to other relevant Webpage and Websites [39]. It is cited as one of the dimensions of Website usefulness [22]. While according to TAM [17], usefulness features of a Website lead into the acceptance and usage of the system by users, it can be concluded that navigability characteristics of a Website are important in influencing customers' intention to purchase from the Website. Therefore, it is suggested that:

H3: Navigation features of a Website will positively influence customers' purchase intention from the Website.

5 Security

It is argued that Website interface features are not enough to attract customers purchase from the Website and other features such as security features are also important [8,10]. According to Chellappa and Pavlou [12], perceived security is "the subjective probability with which consumers believe that their personal information will not be viewed, stored or manipulated during transit or storage by inappropriate parties, in a manner consistent with their confident expectations" (p 359). Security issues are mentioned as one of the main challenges in the adoption of online transactions [4, 44].

In a study by Cyr and Larios [15] on four different nations, it was found that there are significant differences between the four nations' perception on security issues. Belanger et al. [7] used four factors third party privacy seals, third party security seals, privacy statements and website security features to measure customers' intention to purchase from a Website and found that customers value the security features more significantly compared to privacy seals, security seals, and privacy statements. Unless customers are ensured about the security of their information provided to a Website, they do not tend to involve in a transaction with the Website. Therefore, it is suggested that:

H4: Customers' perceived security has a positive influence on their intention to purchase from a Website.

6 Research Methods

This study used closed-ended questions adopted from reliable published sources, shown in the Table 2 and collected data from the respondents in two developing Asian countries Iran and Malaysia with highly different e-readiness indices shown in the Table 1.

Table 1. E-readiness Scores [20: 24]

Country	Overall Score	Connectivity and Technology Infrastructure	Business Environment	Social and Cultural Environment	Legal Environment	Government Policy and Vision	Consumer and Business Adoption	2009 Rank	2008 Rank
Malaysia	5.87	4.8	6.81	5.57	7	6.05	5.8	38	34
Iran	3.43	3.5	4.22	5.23	3	2.65	2.48	68	70

The data was collected from 165 Iranians and 160 Malaysians, witch make a total sample size of 325 respondents. During the data collection, respondents were asked if they are either the citizens or the permanent residents of the country where the data was being collected to ensure that only Malaysians and Iranians or their permanent residents are participated in the data collection. Moreover, the respondents with no online purchase experience were eliminated from the process to ensure that the respondents have the knowledge and experience of purchasing online.

The respondents were asked to fill out the questionnaire with a familiar Website in mind and no specific Website was mentioned for their reference. This method eliminates the influence of Website brand and product/service type from their evaluations [33]. Moreover, since this study collects data from two different nations with different e-readiness rankings, language and culture, asking the respondents to fill out the questionnaire based on their experiences with one internationally famous Website may not be relevant since they may not use the same Website and each nation may have more tendency and experience of using their local Websites rather than common internationally famous Websites.

Before performing any analysis on the data, the data was tested for any outlier or data entry mistakes and corrective actions were taken accordingly. Moreover, the variables were tested for skewness and kurtosis to ensure the normality of the distribution and no problem was found on this regard.

The reliability of the scale was administered by Cronbach's alpha and as shown in the Table 2, was found reliable for all the samples.

Table 2. Scale reliability

Variables	Source		Sample Reliability			No. of Items
	Source	Reliability	Iran	Malaysia	Total	
Security	[25]	0.95	0.823	0.873	0.854	5
Purchase Intention	[34]	0.97	0.864	0.903	0.892	4
Ease of Use	[9]	0.81	0.749	0.761	0.754	3
Navigation	[21]	0.83	0.789	0.833	0.817	3
InformationPresentation	[21]	0.87	0.873	0.777	0.831	3

It was observed that that majority of the Iranian respondents (74.5%) and Malaysian respondents (64.4%) were females. Moreover, it was found that 49.7% of Iranians and 58.1% of Malaysians have the experience of more than one time online purchase while the remaining have the experience of one time online purchase. Moreover, the mean values for the online purchase experience of the samples were 1.75 and 1.97 for Iranians and Malaysians respectively. It indicates that Malaysians on average have more experience of online purchase compared to Iranians.

7 Analysis and Findings

To meet the objectives, different tests including independent sample t test, VIF technique, correlation analysis and multiple linear regression were performed.

7.1 t Test

The descriptive statistics of the two samples showed that that the Iranian respondents have larger mean values and smaller standard deviations compared to the Malaysians in the variable "security". Unlikely, for the other variables, i.e. navigation, information presentation, ease of use and purchase intention, Malaysians have greater mean values than Iranians. In other words, the descriptive analysis show that while Iranians have higher mean values in the security features of a Website, Malaysians have greater mean values in the Website interface elements and purchase intention from the Website.

However, it is natural to have different means between any two samples. Therefore, the independent sample t test with a α level = 0.05 was executed to find if this difference is accidental and is a result of sampling error or is a real difference. Since this study does not predict any direction for the treatment effect, and assumes that the mean values for both samples are approximately equal, a two tailed test was performed.

The results of the 2-tailed independent sample t test showed that the differences between the mean values of the two samples on the variables "security" ($P<0.05$, t-value =0.97, df. =284.26, 2-tailed sig. =0.327) is real and significant. However, the analysis found that the differences between the mean values of the two nations on the variables "navigation" ($P>0.05$, t-value= -2.84, df.=323, 2-tailed sig.=0.005), "information presentation" ($P>0.05$, t-value= -1.96, df=323, 2-tailed sig.=0.05), "purchase intention" ($P<0.05$, t-value= - 4.32, df.= -4.32, df.=310.52, 2-tailed sig.=0.000) and "ease

of use" (P>0.53, t-value= -6.11, df. =322.95, 2-tailed sig.= 0.000) is not real and the variance is not significant.

7.2 Regression Analysis

The multicollinearity was tested using VIF method for all the three samples and no multicolinearity problem was observed in all the three data sets. Then the correlation test was performed for all the three data sets (Total, Iran, Malaysia) and the correlations between all the variables were found significant at the level of 0.01. In order to test the relationships between the independent variables and the dependent variable and compare the results, the linear multiple regression test was performed for the data collected from Iranians, Malaysians and the total data set separately.

It was found that the purchase intention among Iranians is significantly and positively influenced by their perceptions of the security (sig. =0.00, t =3.39), information presentation (sig. =0.00, t =2.87) and ease of use (sig. =0.00, t =5.85) features of the Website while no significant relation was found between Iranians' purchase intention and the navigation features (sig. =0.49, t =0.69) of the Website. Similarly, the results of the analysis for the total data set showed that customers' purchase intention from a Website is positively and significantly influenced by their perceptions toward the security (sig. =0.00, t =4.42), information presentation (sig. =0.00, t =4.93) and ease of use (sig. =0.00, t =5.57) features of the Website while it was found that the navigation features (sig. =0.44, t =0.748) of the Website has no significant influence on customers' purchase intention from the Website. However, it was found that the security (sig. =0.00, t =3.35) and information presentation (sig. =0.00, t =4.43) features of a Website have a positive significant influence on the Malaysians' purchase intention while the ease of use (sig. =0.68, t =0.41) and navigation (sig. =0.35, t =0.93) features of the Website have no significant influence on their purchase intention from the Website.

8 Conclusion and Discussion

This research intends to study the differences in the perceptions of different nations toward a Website features and how these differences influence their purchase intention from the Website. To meet the research objectives, after data screening process, different analytical techniques including independent sample t test, correlation and multiple regression analysis were performed. The results of the t test analysis revealed that Iranians have a greater mean value on the variable "security" features of a Website while the difference in the mean values is real and is not due to sampling errors. However, the t test results showed that there is no real difference between the mean values of Iranians and Malaysians in the other variables "navigation", "ease of use", "information presentation" and "purchase intention" and the differences found in the descriptive analysis are due to sampling error or the natural difference between any given data sets.

These results indicate that the Iranians have more trust in the security mechanisms of the Website they have purchased from since the mean value of their responses is greater than the Malaysians. In other words, the Malaysians have more security concerns while purchasing online. The findings however indicate that there is no real

difference in the perceptions of both the nations on the variables related to Website interface design and purchase intention.

The results of regression analysis indicate that while the Malaysians' purchase intention is significantly and positively influenced by the security and information presentation features of a Website, Iranians' purchase intention is significantly influenced by security, information presentation and ease of use features of a Website. The analysis results of the total data set shows similar results with Iranians. These findings suggest that the significance of the relations exist in an e-commerce related model may differ from a nation to another, which may be the result of their e-readiness, culture or any other differences existing between nations. These findings raise questions on the generalization power of the findings of researchers in one country, which may not hold true in another country. In other words, while a research model and the relations among its variables may be supported in one country, same model and variables may not be supported in another country; which reduces the generalization power of the findings of many studies due to their context-sensitive nature. The results also suggest that companies marketing internationally and addressing different nations through their e-commerce strategies and technologies, should not formulate and implement a single global e-commerce strategy and their strategies should be tailored for each region and nation.

9 Limitations and Future Studies

This study is not free from limitations. First, the study collected data from the respondents using a non probability sampling technique based on convenience which eliminate the generalization of the findings as they represent a particular portion of their society. The second limitation is that the study focused only on a few determinant factors of purchase intention and a more comprehensive model with more variables should be studied to have a clearer picture of the differences might exist among the nations. Furthermore that the data was collected from those who have experienced of at least one time online purchase and the data does not represent the perceptions of all the citizens of a nations, but only those who have experienced online purchase, while others, who have not involved any online transaction may have different perceptions, which have prevented them from involving in such transactions. Furthermore, the study collected data from two Asian developing countries. In order to investigate if there are serious differences between developed and developing nations, future studies should collect data from both developed and developing countries and compare the results to find any probable differences either in their perceptions or in the significant and direction of the relations among the variables of the research model.

References

1. Abbasi, A.: E-Commerce Development in Iran. Webology 4(4), Article 49 (2007), http://www.webology.ir/2007/v4n4/a49.html
2. Ajzen, I., Fishbein, M.: Understanding Attitudes and Predicting Social Behavior. Prentice-Hall, Englewood Cliffs (1980)

3. Ajzen, I.: From Intentions to Actions: a Theory of Planned Behavior. In: Kuhl, J., Becaman, J. (eds.) Action-control: From Cognition to Behavior, pp. 11–39. Springer, Heidelberg (1985)
4. Aladwani, A.M.: Online Banking: a Field Study of Drivers, Development Challenges, and Expectations. International Journal of Information Management 21(4), 213–225 (2001)
5. Alcaniz, E.B., Ruiz-Mafé, C., Alda´s-Manzano, J., Sanz-Blas, S.: Influence of Online Shopping Information Dependency and Innovativeness on Internet Shopping Adoption. Online Information Review 32(5), 653–667 (2008)
6. Aldas-Manzano, J., Lassala-Navarre, C., Ruiz-Mafé, C., Sanz-Blas, S.: Key drivers of Internet Banking Services Use. Online Information Review 33(4), 672–695 (2009)
7. Belanger, F., Hiller, J.S., Smith, W.J.: Trustworthiness in Electronic Commerce: the Role of Privacy, Security, and Site Attributes. Journal of Strategic Information Systems 11, 245–270 (2002)
8. Berthon, P., Pitt, L., Cyr, D., Campbell, C.: E-readiness and Trust: Macro and Micro Dualities for E-commerce in a Global Environment. International Marketing Review 25(6), 700–714 (2008)
9. Chang, H.H., Chen, S.W.: Consumer Perception of Interface Quality, Security, and Loyalty in Electronic Commerce. Information & Management 46, 411–417 (2009)
10. Chang, H.H., Chen, S.W.: The Impact of Online Store Environment Cues on Purchase Intention TRUST and Perceived Risk as a Mediator. Online Information Review 32(6), 818–841 (2008)
11. Chau, P.Y.K., Au, G., Tam, K.Y.: Impact of Information Presentation Modes on Online Shopping: an Empirical Evaluation of Broadband Interactive Shopping Service. Journal of Organizational computing and electronic commerce 10(1), 1–22 (2000)
12. Chellappa, R.K., Pavlou, P.: Perceived Information Security, Financial Liability and Consumer Trust in Electronic Commerce Transactions. Logistics Information Management 15(5), 358–368 (2002)
13. Chen, Y.-H., Barnes, S.: Initial Trust and Online Buyer Behaviour. Industrial Management & Data Systems 107(1), 21–36 (2007)
14. Cyr, D., Bonanni, C., Ilsever, J.: Design and E-loyalty Across Cultures in Electronic Commerce. In: Sixth International Conference on Electronic Commerce. ACM, New York (2004), 1-58113-930-6/04/10
15. Cyr, D., Larios, H.: Managing E-loyalty through Experience Design: Results of a Four Nation Study (2004), http://www.diannecyr.com/docs/managing_eloyalty.pdf
16. Cyr, D., Headand, M., Larios, H.: Colour Appeal in Website Design Within and Across Cultures: A Multi-method Evaluation. International Journal of Human–Computer Studies (2009) (publication under process, accessed on September 9, 2009), doi:10.1016/j.ijhcs.2009.08.005
17. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 13(3), 319–340 (1989)
18. DeLone, W.H., McLean, E.R.: The Delone and Mclean Model of Information Systems Success: a Ten-year Update. Journal of Information Systems Management 19(4), 9–26 (2003)
19. DeLone, W.H., McLean, E.R.: Information Systems Success: the Quest for the Dependent Variable. Information Systems Research 3(1), 60–95 (1992)
20. Economist Intelligence Rankings (2009), http://www-935.ibm.com/services/us/gbs/bus/pdf/e-readiness_rankings_june_2009_final_web.pdf

21. Éthier, J., Hadaya, P., Talbot, J., Cadieux, J.: Interface Design and Emotions Experienced on B2C Web sites: Empirical testing of a research model. Computers in Human Behavior 24, 2771–2791 (2008)
22. Fang, X., Holsapple, C.W.: An Empirical Study of Web Site Navigation Structures' Impacts on Web Site Usability. Decision Support Systems 43, 476–491 (2007)
23. Fishbein, M., Ajzen, I.: Belief, Attitude, Intention and Behavior: an Introduction to Theory and Research. Addison-Wesley, Reading (1975)
24. Gurung, A.: Empirical Investigation of the Relationship of Privacy, Security and Trust with Behavioral Intention to Transact in E-commerce, unpublished PhD dissertation, The University of Texas at Arlington, US (2006)
25. Ha, S., Stoel, L.: Consumer E-shopping Acceptance: Antecedents in a Technology Acceptance Model. Journal of Business Research 62, 565–571 (2009)
26. Keat, T.K., Mohan, A.: Integration of TAM Based Electronic Commerce Models for Trust. The Journal of American Academy of Business, Cambridge, 404–410 (September 2004)
27. Kim, C., Tao, W., Shin, N., Kim, K.: An Empirical Study of Customers' Perceptions of Security and Trust in E-payment Systems. Electronic Commerce Research and Applications 9, 84–95 (2010)
28. Kim, D.-Y., Lehto, X.Y., Morrison, A.M.: Gender Differences in Online Travel Information Search: Implications for Marketing Communications on the Internet. Tourism Management 28, 423–433 (2007)
29. Kwon, O.B., Kim, C., Lee, E.J.: Impact of Website Information Design Factors on Consumer Ratings of Web-based Auction Sites. Behaviour & Information Technology 21(6), 387–402 (2002)
30. Loiacono, E.T., Watson, R.T., Goodhue, D.L.: WebQual: A Measure of Web site Quality. In: Marketing Educators' Conference: Marketing Theory and Applications, vol. 13, pp. 432–437 (2002)
31. Monsuwe, T.P., Dellaert, B.G.C., Ruyter, K.: What Drives Consumers to Shop Online? A literature review. International Journal of Service Industry Management 15(1), 102–121 (2004)
32. Moustakis, V., Tsironis, L., Litos, C.: A Model of Website Quality Assessment. The Quality Management Journal 13(2), 22–37 (2006)
33. Peikari, H.R.: Website Design, Security, Privacy, Verification and Customer's Trust in e-Commerce: An Empirical Examination. In: Proceedings of 3rd Asia Pacific Marketing Conference 2009, Kuching, Malaysia, pp. 9 – 11 (December 2009)
34. Poddar, A., Donthu, N., Wei, Y.: Web Site Customer Orientations, Web Site Quality, and Purchase Intentions: The Role of Web Site Personality. Journal of Business Research 62, 441–450 (2009)
35. Roca, J.C., Garcia, J.J., Vega, J.J.: The Importance of Perceived Trust, Security and Privacy in Online Trading Systems. Information Management & Computer Security 17(2), 96–113 (2009)
36. Rousea, S.V., Haas, H.A.: Exploring the Accuracies and Inaccuracies of Personality Perception Following Internet-mediated Communication. Journal of Research in Personality 37, 446–467 (2003)
37. Shama, A.: An Empirical Study of the International Marketing Strategies of E-commerce Companies. Thunderbird International Business Review 47(6), 695–709 (2005)
38. Tarafdar, M., Zhang, J.: Analysis of Critical Website Characteristics: a Cross-category Study of Successful Websites. Journal of Computer Information Systems, 14–24 (Winter 2005-2006)

39. Von Dran, G.M., Zhang, P., Small, R.: Quality Websites: an Application of the Kano Model to Website Design. In: Proceedings of the Americas Conference on Information Systems, Dallas, Texas, August 7-8 (2002)
40. Wang, C.C., Yang, H.W.: Passion for Online Shopping: the Influence of Personality and Compulsive Buying. Social Behavior and Personality 36(5), 693–706 (2008)
41. Wen, I.: Factors Affecting the Online Travel Buying Decision: a Review. International Journal of Contemporary Hospitality Management 21(6), 752–765 (2009)
42. Wu, I., Chen, J.: An extension of Trust and TAM Model with TPB in the Initial Adoption of On-line Tax: An Empirical Study. International Journal of Human-Computer Studies 62, 784–808 (2005)
43. Yang, Z., Jun, M., Peterson, R.T.: Measuring Customer Perceived Online Service Quality Scale Development and Managerial Implications. International Journal of Operations & Production Management 24(11), 1149–1174 (2004)
44. Yousafzai, S.Y., Pallister, G.H., Foxall, G.R.: A Proposed Model of E-trust for Electronic Banking. Technovation 23, 847–860 (2003)

Deployment of ERP Systems at Automotive Industries, Security Inspection (Case Study: IRAN KHODRO Automotive Company)

Hatamirad Ali[1] and Mehrjerdi Hasan[2]

[1] IT/IS department, IRAN KHODRO Automotive Company and Department of E-Commerce, NooreTouba Universsity, Tehran, Iran
a.hatamirad@ikco.com
[2] Department of electrical enginering, Université du Québec, Montreal, Canada
hasan.mehrjerdi.1@ens.etsmtl.ca

Abstract. Automotive industry and car production process is one of the most complex and large-scale production processes. Today, information technology (IT) and ERP systems incorporates a large portion of production processes. Without any integrated systems such as ERP, the production and supply chain processes will be tangled. The ERP systems, that are last generation of MRP systems, make produce and sale processes of these industries easier and this is the major factor of development of these industries anyhow. Today many of large-scale companies are developing and deploying the ERP systems. The ERP systems facilitate many of organization processes and make organization to increase efficiency. The security is a very important part of the ERP strategy at the organization, Security at the ERP systems, because of integrity and extensive, is more important of local and legacy systems. Disregarding of this point can play a giant role at success or failure of this kind of systems. The IRANKHODRO is the biggest automotive factory in the Middle East with an annual production over 600.000 cars. This paper presents ERP security deployment experience at the "IRANKHODRO Company". Recently, by launching ERP systems, it moved a big step toward more developments.

Keywords: ERP, security, automotive industry, integrated systems.

1 Introduction

The evolution of information technology systems from the beginning was quite similar in all industries and activity areas. In the 1960s and 1970s companies chose a hardware provider, and from there some basic software development products (programming languages), and started to develop their business applications. Most companies started with critical areas, like accounting and financial applications, that were somehow easier. Later, these companies advanced and introduced applications in other, more complex areas like distribution and production. In any case, they always made their own development using the previously chosen hardware and software. Already in the 1970s there were some companies that realized the possibility of developing business software that could be used by different companies; the opportunity existed to develop the applications only once and then sell the software to other companies. The term "Enterprise

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 160–169, 2010.

Resource Planning" originally derived from manufacturing resource planning (MRP II) that followed material requirements planning (MRP). [1]

ERP systems typically handle the manufacturing, logistics, distribution, inventory, shipping, invoicing, and accounting for a company. ERP software can aid in the control of many business activities, including sales, marketing, delivery, billing, production, inventory management, quality management, and human resource management. ERP systems saw a large boost in sales in the 1990s as companies faced the Y2K problem (real or imagined) in their "legacy" systems. Many companies took this opportunity to replace such information systems with ERP systems. This rapid growth in sales was followed by a slump in 1999, at which time most companies had already implemented their Y2K solution. [2]

ERP systems are often incorrectly called back office systems indicating that customers and the general public are not directly involved. This is contrasted with front office systems like customer relationship management (CRM) systems that deal directly with the customers, or the E-Business systems such as E-Commerce, E-Government, E-Telecom, and E-Finance, or supplier relationship management (SRM) systems. ERP systems are cross-functional and enterprise-wide. All functional departments that are involved in operations or production are integrated in one system. In addition to areas such as automotive manufacturing, warehousing, logistics, and information technology, this typically includes accounting, human resources, marketing and strategic management.

ERP II, a term coined in the early 2000s, is often used to describe what would be the next generation of ERP software. This new generation of software is web-based and allows both employees and external resources (such as suppliers and customers) real-time access to the system's data, especially at automotive companies with several partners and supplier producers.

EAS (Enterprise Application Suite) is a new name for formerly developed ERP systems which include –almost- all segments of business using ordinary Internet browsers as thin clients. Though traditionally ERP packages have been on-premise installations, ERP systems are now also available as Software as a Service.

Security at the ERP systems, because of integrity and extensive, is more important of local and legacy systems. To realization of this goal, gradually the legacy systems at organizations will replace to integrated systems, such as ERP. Today ERP systems are an undeniable part of enterprise organizations such as automotive industries.

2 ERP Systems

Enterprise Resource Planning systems (ERP) are pre engineered packages designed for internal information processing and facilitating integration of the information of the firms and fast responding to the customers. At another word, ERP is an integrated computer-based system used to manage internal and external resources including tangible assets, financial resources, materials, and human resources. It is a software architecture whose purpose is to facilitate the flow of information between all business functions inside the boundaries of the organization and manage the connections to outside stakeholders. Built on a centralized database and normally utilizing a common computing platform, ERP systems consolidate all business operations into a uniform and enterprise wide system environment. [3]

Implementation of these systems is always time consuming and expensive and all parts of the organization will be challenged with the implementation process. Evidence indicates that despite the excessive time and cost, unfortunately most firms haven't implemented these systems successfully and others which climes success, have some problems with after implementation. For protecting the organizations against the risks and dangers of ERP, it is necessary that managers arrange an investigation for organizational readiness assessment before ERP implementation. [4]

At figure 1, the ERP system components displayed, such as financial management, customer relationship management, supply chain management, human resource management and manufacturing resource management.

Fig. 1. ERP system components

3 ERP History

3.1 Before 1960

During this period, software such as Processors BOM was deployed. The principal goal of this software was to explore of required materials to produce materials. This software did not consider volume and dimension of productions or in other word Lot Sizing as well delivery time.

3.2 Between 1960 to 1970

At the beginning of this decade the software's focus was over the stock control systems. Although, at the end of this decade, the MRP concept or material planning was defined and the MRPI software was deployed by IBM.

Execution on the expensive mainframes which usually used on the university or military services was the main problem of this software and similar ones.

3.3 Between 1970 to 1980

In this duration, the attention was on the MRP I with focus on the total production planning or MPS as well MRP development. In these systems, failure possibility at supply chain reduced to minimum. However, in these systems effort were done only on work planning and did not support any other production sources.

3.4 Between 1980 to 1990

In this duration, the MRP II development was occurred that support many of production sources. In the first years of this decade, the DPR or distributed planning systems that operate independent of MRP II were mixed to these systems. Moreover, the communication problem between these two systems was solved.

3.5 Between 1990 to 2000

In this period, MRP II was developing and support systems (DSS) were added to MRPII. At 1995, internet came into ERP systems and between1998 to 2000 EDI and ERP were combined. Today's ERP II systems are developing based on web and has more concentration on SCM support.

4 ERP Advantages

In the absence of an ERP system, a large manufacturer may find itself with many software applications that cannot communicate or interface effectively with one another. Tasks that need to interface with one another may involve:

1. ERP systems connect the necessary software in order for accurate forecasting to be done. This allows inventory levels to be kept at maximum efficiency and the company to be more profitable.
2. Integration among different functional areas to ensure proper communication, productivity and efficiency
3. Design engineering (how to best make the product)
4. Order tracking, from acceptance through fulfillment
5. The revenue cycle, from invoice through cash receipt
6. Managing inter-dependencies of complex processes bill of materials
7. Tracking the three-way match between purchase orders (what was ordered), inventory receipts (what arrived), and costing (what the vendor invoiced)
8. The accounting for all of these tasks: tracking the revenue, cost and profit at a granular level.

ERP Systems centralize the data in one place. Benefits of this include:

1. Eliminates the problem of synchronizing changes between multiple systems consolidation of finance, marketing and sales, human resource, and manufacturing applications
2. Permits control of business processes that cross functional boundaries

3. Provides top-down view of the enterprise (no "islands of information"), real time information is available to management anywhere, anytime to make proper decisions.
4. Reduces the risk of loss of sensitive data by consolidating multiple permissions and security models into a single structure.
5. Shorten production lead-time and delivery time
6. Facilitating business learning, empowering, and building common visions

Some security features are included within an ERP system to protect against both outsider crime, such as industrial espionage, and insider crime, such as embezzlement. A data-tampering scenario, for example, might involve a disgruntled employee intentionally modifying prices to below-the-breakeven point in order to attempt to interfere with the company's profit or other sabotage. ERP systems typically provide functionality for implementing internal controls to prevent actions of this kind. ERP vendors are also moving toward better integration with other kinds of information security tools. [5]

5 ERP Disadvantages

Problems with ERP systems are mainly due to inadequate investment in ongoing training for the involved IT personnel - including those implementing and testing changes - as well as a lack of corporate policy protecting the integrity of the data in the ERP systems and the ways in which it is used.
Disadvantages:

1. Customization of the ERP software is limited.
2. Re-engineering of business processes to fit the "industry standard" prescribed by the ERP system may lead to a loss of competitive advantage.
3. ERP systems can be very expensive (This has led to a new category of "ERP light" solutions)
4. ERPs are often seen as too rigid and too difficult to adapt to the specific workflow and business process of some companies (this is cited as one of the main causes of their failure).
5. Many of the integrated links need high accuracy in other applications to work effectively. A company can achieve minimum standards, and then over time "dirty data" will reduce the reliability of some applications.
6. Once a system is established, switching costs are very high for any one of the partners (reducing flexibility and strategic control at the corporate level).
7. The blurring of company boundaries can cause problems in accountability, lines of responsibility, and employee morale.
8. Resistance in sharing sensitive internal information between departments can reduce the effectiveness of the software.
9. Some large organizations may have multiple departments with separate, independent resources, missions, chains-of-command, etc, and consolidation into a single enterprise may yield limited benefits.

6 ERP Deployment at IRAN KHODRO

The SAP Company, one of the biggest integrated systems providers, is in cooperation with IRAN KHODRO to deploy ERP systems since 2005.

The goal of deploying these systems is reengineering of production and supply chain processes, standardize and integrity of database and application systems.

7 Security at ERP Systems

Security is increasingly being considered one of the key points to boost electronic commerce over the Web. The ERP systems have always established security as one of the critical topics both for the implementation and correct deployment of integration Solutions. Every professional involved in modern ERP systems is aware that leveraging security technology and measures and a sound security policy is mandatory. The information stored in the systems we support ranks among a company's most important and valuable assets. Moreover, addressing security during and after an ERP implementation not only protects valuable business information; it ensures continuous and stable systems operations.

Security can be defined from two different perspectives that have in common the objective of protecting the company systems and information assets. These two perspectives are as follows:

1. *Security* as the protection measures and policies against unauthorized accesses by illegitimate users (both internal and external).
2. *Security* as protection measures against hardware, software, or any other type of environmental failures (disasters, fires, earthquakes, and others) using.

7.1 Security Policies at ERP

Companies must implement some type of security policy to protect their assets, but also they are required to comply with their country's legal obligations, business agreements, and industry laws and regulations. For instance, many countries have some forms of for protecting confidential data of employees. It is also very important to keep all financial records for tax authorities. And in terms of business partners, it is of great Importance to ensure the confidentiality of commercial agreements with vendors or customers.

Modern information systems and technologies are both the means and the containers the strategic and operative business information. They are the known but hidden treasures of companies, and companies need to keep their treasures secure.

The *Security Policy* is the set of procedures, standards, roles, and responsibilities covering and specifying all the security and organizational measures that companies follow to protect their business from threats and vulnerabilities. An approach to security will have the objective of building a strong security policy and should start by assessing risk analysis to implement, monitor, and enforce such policy. It is important to realize that security implementation never ends and must be continually updated, reviewed, communicated, implemented, monitored, and enforced.

It's mandatory the ERP systems have these features at security:

1. Set up private communication channels.
2. Use strong authentication mechanisms.
3. Implement group concept in Java.
4. Provide evidence of business transactions.
5. Enforce auditing and logging.

Among these objectives these security services have to available for ERP too:

1. The use of client and server certificates for user *authentication*
2. Single Sign-On solutions to access the full range of ERP components and solution
3. The *role-based* concept, which involves activity groups and authorizations
4. Deployment of firewalls between systems and networks, as well as secure
5. protocols such as HTTPS (HTTP over SSL)
6. SNC (Secure Network Communications) and SSF (Secure Store and Forward) for Compliance with security standards.

7.2 Basic Security Processes at ERP Systems

These basic security processes are the standards and in the all ERP systems make the Security based:

7.2.1 Authentication

Authentication is the process that is used for verifying that users, programs, or services are actually who they say they are. Authentication is the cornerstone of any security infrastructure or technology.

The ERP's standard User Authentication verifies a user's identity through the use of logon passwords. (Unsuccessful logon attempts will cause the session to terminate and activate user locks.) As standard security measures, have to provide several login profile parameters and an initial set of password rules that expand on according to your needs. Limitations on ERP standard authentication pertain to the legal export rules of different countries regarding encryption software and algorithms.

7.2.2 Authorization

Authorization is the process that is used for determining what accesses or privileges are allowed for users. Authorizations are enforced by means of access controls, which are in charge of restricting user accesses.

In the all ERP systems, standard User Authorization secures user access to business data and transactions, ensuring that only preauthorized users gain access to data and processes.

7.2.3 Privacy

Privacy is the process that can be used for ensuring that data or information sent over a network or communication line is not accessed or read by unauthorized persons. A usual way of granting privacy is by using *cryptography* technology. Both authorization and privacy ensure the confidentiality of data and information.

7.2.4 Integrity

Integrity is the process that verifies that nothing or nobody modifies data from a source to a target. *Integrity* can be enforced by means of digital signatures, digital envelopes, and the use of the SNC and SSF components.

7.2.5 Proof of Obligation

Obligation or proof of obligation is necessary for confirming and guaranteeing that a business message is correct so it can be considered a business transaction between business partners. For this reason in electronic commerce there must be enough security mechanisms to guarantee the *no repudiation* of business messages.

7.2.6 Auditing

Auditing is the process of collecting and analyzing security data for verifying that the security policy and rules are complied with. *Accounting* is a way of measuring and/or restricting the use of system resources and as such is a form of authorization.

7.2.7 Cryptography

Cryptography is the technique based on mathematical algorithms and other methods to encode data and thus prevent data from being read or disclosed. Cryptography is commonly defined as the science of secret writing.

7.2.8 Public Key Cryptography

Public key cryptography is based on mathematical functions of one direction, meaning that it is impossible to observe the results. With this type of system each user that originates communications or messages has two keys:

1. A private one (secret)
2. A public one that is distributed to their communication partners

Every message that is sent with public key can only be decrypted using the private key.

7.2.9 Digital Signature

Digital signatures are special appendixes that are added to the digital documents to show the authenticity of the origin and the integrity of those documents. A digital signature is equivalent to the traditional hand-written signatures on paper documents. When someone tries to modify a handwritten signature illegally, there are usually clues that can be detected by physical means. This is usually what guarantees the authenticity and integrity of data and information contained.

The digital signature must guarantee the same elements although using technological means. The first important point is that each digital signature will be different in every document. Otherwise it could be easy to copy and falsify digital signatures. For this reason the digital signature will depend on the document that is being signed using a mathematical function. This mathematical relationship allows for later verification of the validity and authenticity of the document.

7.2.10 Secure Socket Layer Protocol (SSL)

HTTP is the default protocol for transferring files on the World Wide Web. HTTP transports Web sites as plain-text files. So it is possible that a third party having access to the network can read or alter the data sent. The protocol has no proper mechanisms to ensure authentication and confidentiality for the data. For that purpose SSL encryption can be used. The HTTPS protocol transfers HTTP over an SSL connection. HTTPS offers options to encrypt the data and to identify the other party by its digital certificate.

7.2.11 Single Sign-On (SSO)

The Single Sign-On solution is a standard; users only need to enter their passwords once when they initially log on to the security system or the operating system.

8 Conclusion

Enterprise Resource Planner (commonly known as ERP) software is a concept that started in the 1970s and was meant to provide computerized solutions for integrating and automating business processes across companies' back offices, such as the production, financial, supply chain and logistics, or human resources departments. The idea behind ERP was that companies could see a cost reduction and better efficiency in the way they operated with their business partners (customers, providers such as part supplier providers, banks, authorities, etc.) and also in the way their users could access and process the information.

From that concept, there were already several solutions in the market during the 1980s and beginning of the 1990s. The adoption of ERP software revolutionized the way companies conduct their traditional business.

Modern information systems and technologies are both the means and the containers of the strategic and operative business information. They are the known but hidden treasures of companies, and companies need to keep their treasures secure.

The Security Policy is the set of procedures, standards, roles, and responsibilities covering and specifying all the security and organizational measures that companies must follow to protect their business from threats and vulnerabilities. An approach to security will have the objective of building a strong security policy and should start by assessing a risk analysis to implement, monitor, and enforce such policy. It is important to realize that security implementation never ends and must be continually updated, reviewed, communicated, implemented, monitored, and enforced.

The security strategy and risk analysis must first consider these basic issues:

1. What is to be protected?
   Companies must identify those assets -such as critical information (production data, supply chain data, customer list, employee personal data, contracts), hardware, software, intangibles (hours of operation, cost of on revenue, non-production) or others- that require some type and some degree of protection against unwanted and unauthorized access, which could damage or destroy to some degree such assets.
   a. *Which are the possible threats?*

The second security issue is to identify the possible sources of attack and the degree of vulnerability of infrastructure. Threats are of different type and nature and sometimes unknown. They are often intentional, but can also be unintentional. They can be external threats or can be internal (for instance, by other geographical locations or by burned-out or frustrated employees).

2. What protection measures can be taken?
   Finally, the risk analysis and the security policy must identify the best security measures to implement and enforce such policy efficiently. Measures can be standard measures included in the information system capabilities, additional and external security infrastructure, and behavioral rules. For instance, a basic and strong security measure is the password that users must provide to access systems; however, it is almost impossible with technical means to know whether someone told his or her password to someone else.

Efficiency in security policy means that measures do not include awkward procedures that would obstruct or make users' jobs more difficult. Security policies always follow a principle of controls, which means that the security strategy must approach the balance between risks and control measures.

As indicated, security is a continuous process due to the fact that new assets, new threats, or new technology can be identified as well as some threats or assets that are obsolete and no longer need protection. These facts will make the security policy a living entity that also includes the retraining of employees.

At the ERP systems, you can improve security by

1. Designing and implementing a secure systems infrastructure by means of firewalls and setting password policies and parameters
2. Setting the most appropriate values for security-related instance profile parameters
3. Using external security products
4. Establishing a security policy and efficiently communicating it
5. Creating a security checklist that can be periodically tested either manually or automatically so you can evaluate the efficiency of your security policy
6. Enforcing the security policy by means of logging and auditing
7. Monitoring security alerts and locating threats
8. Establishing a procedure for constant update of the security policies

References

[1] Anderegg, T.: MRP/MRPII/ERP/ERM - Confusting Terms and Definitions for a Murkey Alphabet Soup
[2] Monk, E., Wagner, B.: Concepts in Enterprise Resource Planning, 2nd edn. Thomson Course Technology, Boston (2006) ISBN 0619216638
[3] Bidgoli, H.: The Internet Encyclopedia, vol. 1, p. 707. John Wiley & Sons, Inc., Chichester (2004)
[4] Saremi, Khani, M., Abedi: Khnowledge management magazine, Tehran, Iran, vol. 77 (2008)
[5] Walsh, K.: The ERP Security Challenge. CSOonline. CXO Media In (2008)

Governance and Risk Management of Network and Information Security: The Role of Public Private Partnerships in Managing the Existing and Emerging Risks

Jyoti Navare and Orhan Gemikonakli

Middlesex University, London
{j.navare,o.gemikonakli}@mdx.ac.uk

Abstract. Globalisation and new technology has opened the gates to more security risks. As the strategic importance of communication networks and information increased, threats to the security and safety of communication infrastructures, as well as information stored in and/or transmitted increased significantly. The development of the self replicating programmes has become a nightmare for Internet users. Leading companies, strategic organisations were not immune to attacks; they were also "hacked" and overtaken by intruders. Incidents of recent years have also shown that national/regional crisis may also trigger cyber attacks at large scale. Experts forecast that cyber wars are likely to take the stage as tension mounts between developed societies. New risks such as cyber-attacks, network terrorism and disintegration of traditional infrastructures has somewhat blurred the boundaries of operation and control. This paper seeks to consider the risk management and governance and looking more specifically at implications for emerging economies.

1 Introduction

Globalisation and new technology has opened the gates to more security risks. As the strategic importance of communication networks and information increased, threats to the security and safety of communication infrastructures, as well as information stored in networked environments and/or transmitted increased significantly. 38 years after the development of the first self replicating programme, a variety of illegitimate software developments followed; viruses, worms, Trojan horses, zombies and so on. Leading companies, strategic organisations were not immune to attacks; they were also "hacked" and overtaken by intruders. Incidents of recent years have also shown that national/regional crisis may also trigger cyber attacks at large scale; in 2007 Russia was blamed for waves of attacks on Estonian IT systems which continued for three weeks. About a year later, following the increasing tension between South Ossetia and Georgia, Russia was accused of similar actions. Cyber attacks are no longer the act of lone rangers. Experts forecast that cyber wars are likely to take the stage as tension mounts between developed societies. One concern is the cosmopolitan nature of western capitals and the royalty many ethnic minorities living in these cities nourish for their country of origin. As attacks turn into a concerted effort, there is a manifestation of a dynamic

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 170–177, 2010.

world where society, governments, and businesses need to share a collective role in managing existing and emerging complex interdependent security challenges. New risks such as cyber-attacks, network terrorism and disintegration of traditional infrastructures has somewhat blurred the boundaries of operation and control. Traditional approaches can no longer provide effective protection. It is no longer viable to rely on technical personnel's precautions. It is obvious that disaster recovery can only minimise the damage done, pro-active measures can only reduce risks. While disaster recovery procedures are further developed, preventive measures are excelled, it is essential that new approaches to preventing disasters are also explored. This is where the role of transnational organisations can be further explored. Can collaborations, such as between Public and Private Organisations, reduce risks and provide both economic and regulatory safety nets? Can transnational organisations play a role in fighting cyber attacks? What are the mechanisms of fighting the "enemy within"?

There is growing body of research within the field of network governance providing theoretical concepts for public-private collaboration that can be applied for the management of new risks. This paper seeks to consider the risk management and governance and looking more specifically at implications for emerging economies.

2 The Key Risks of Global Networks

The first self replicating programme may have been developed for fun meaning no harm. However, the way it paved has led to the worst nightmare of the cyber world. It is common knowledge that large enterprise, including some in software development suffered attacks. A study carried out by Symantec showed that amongst other risks, cyber attacks play a major role in threatening business security (see Table 1) [1]. According to this report, large enterprise suffers a loss of $2.8 m annually.

Table 1. Most significant risks

Cyber attacks	42%
Traditional criminal activity	17%
Brand-related events	17%
Natural disasters	14%
Terrorism	10%

Although lone hackers would continue admire their hacking skills, the growing concern is the developments towards a cyber war. In other words, battle fields are recreated in the cyber world staging the war of giants. A quick look at some large scale attacks can highlight the importance of the emerging risks: Last December, Twitter, a popular social network was brought to a standstill; this was the third major incident in a year. An organisation calling itself *Iranian Cyber Army* assumed responsibility for this malicious act that crippled a social network worth $1 billion and used by 25 million members. It is

believed that this was in response to the use of Twitter during the Iranian elections that led to widespread protests by the opposition groups. Had Twitter been in the stock market, the results would have been devastating. The same organisation claimed responsibility for another attack; this time the victim was China's most popular search engine, Baidu. In both attacks, DNS entries were altered to direct users to a different website. The attacks to two Iranian sites were assumed to be in response to attack on Baidu; the intruders are not known.

The attacking forces are changing from lone cowboys to groups, groups to well organised – "national" – armies. 2007 attacks on Estonian networked services hit a large number of government and corporate services. These attacks can cripple communication backbones bringing the country's communication to a halt. While Estonian authorities claimed to have traced these attacks to Russia, and suggested that Kremlin was behind these, Russian authorities denied any involvement. During the Ossetia crisis, as a result of waves of attacks, Georgia lost all of its Internet communications for a period of time. Georgian officials blamed Russia for this. War was taking place in the air, on the ground, from the sea and in the cyber space. The situation was made even more difficult by the use of servers from outside Russia, blocking traffic from Russia wouldn't have helped. A detailed account of *The Georgia's Cyberwar* is given in [Rios et. al., 2009a].

China's recent fight with Google is well-known. Chinese army has shown an awareness of the new approaches to war including *commercial, economical,* and *ecological war* as well as an anticipation of the challenges of the Information Warfare for more than a decade. Documents and a book published by army personnel towards the end of the 20th century showed their awareness and strategic approach to this [Rios et. al., 2009b].

The destruction of wars in recent centuries and the two world wars forced nations of the world to the agreement of treaties to limit the damages of wars between nations; Hague Conventions, Geneva Convention, and more than 10 treaties addressing nuclear disarmament/restrictions (starting with Partial Test Ban Treaty in 1963, following with others such as SALT I, SALT II etc. and finally reaching to the New Start Treaty in 2010 – yet to be ratified.) Now that cyber war is threatening the world's peace, urgent action is needed at international level to combat the threat of a cyberwar; a war to which no nation is immune too.

3 Public-Private Partnerships in Reducing Risks and Increasing Governance in the Field of Network Security

We establish that network security although important clearly has significant risks attaching. Furthermore there is risk accumulation in that added investments in enabling and maintaining security can be significant and demonstrating return on security investment is sometimes difficult. Public private solution offer however not only financial but non-financial benefits such as support services in crises – although arguably these could have economic implications. The extension of private finance adds to the public debt and the measuring of importance and value is significant if private financing is not seen to be illusory. Collaborative partnerships are valuable where they bring in strong knowledge, skills, operational know and resources how to solve

cyber security threats and crises. As de Langen (2004) pointed out, the scale, scope and diversity of projects , whether involved in the network security or otherwise provides a need for partnership arrangements between different bodies to provide sufficient funds, to furnish the required range of expertise and to co-ordinate provision.

According to INSA (Intelligence and National Security Alliance US) "An effective partnership has a representative group of members, large enough to be sufficiently inclusive, but small enough to retain the ability to act quickly. A circumscribed role for government with specific tasks and responsibilities laid out clearly. Industry and private groups should take the lead. Properly motivated members with significant interest and stakes connected to the problem [*sic*]"

The function for INSA cyber task group involves "inspection and enforcement of standards upon suppliers and Internet Service Providers (ISPs). Ability to watch networks, searching for and analyzing future threats and warning all users before an emergency occurs. Ability to respond to attacks, through warnings and technical fixes, as well as plan for the recovery of crucial systems after an emergency. Necessary protection for privacy and free speech, individual rights and business concerns, cognizant of government needs. Resulting implementation should work toward collaborative solutions. Mechanism for international collaboration on cyber security [*sic*]" [INSA, 2009]

Some services, such as public security are natural functions of the state, because of the lack of sufficient incentives for private-sector organisations to take action in the wider public interest . Even then, some delegation may be possible under supervision (Chalfin, 2007), with the incentives for proper performance by the private-sector consisting of the need to retain operating licensing and the avoidance of more intrusive supervision (Handley-Schachler and Navare, 2010). There are certain security breaches which may not be in the commercial interests of private companies such as those that pose a national threat. In such cases this is best controlled by the public-sector as there lad organisation – as seen with INSA in the United States.

4 The Robustness of PPPs and the Risks That Can Arise from the Collaboration

As discussed, PPPs enable the combination of skills from public and private-sector partners to provide expertise in a wider variety of functions. They can, in some instances, also enable an improvement in risk allocation, by seeking to allocate the cost of some risks to the party responsible for causing the risk or best able to manage it. One form of PPP is the Private Finance Initiative (PFI), which can involve the provision of assets and services by a private-sector partner who is also responsible for financing the assets or services for direct or indirect use in service provision by a public-sector partner (Handley-Schachler and Navare, 2010)

There are specific risks that arise out of collaborations. Risk identification, allocation and management are vital issues with partnership projects presenting both project and partnership risks as well as opportunities such as risk transference. A number of risks can arise out of partnership arrangements. These risks, however, can be exacerbated by the difficulties in managing the activities and anticipating the decisions of commercial and public-sector partners. At the same time, the creation of partnerships

can provide broader expertise in managing risks and greater financial resources to enable the risk to be borne and shared by the partnership. These fall essentially into four categories: operational, investment, political and supply side. Operational risks are those that affect the physical management of network operations including human safety in this. Investment risk is the probability of a low or nil return on investment and political risks are where national security can result in governments with governments coercing partners to participate in resource provision which can have financial implications. Supply-side risks include costs of design, construction, service provision and service interruptions and connectivity.

It is clear that that PPPs can contribute to risk mitigation, although this is dependable on the the robustness of the arrangement, Ernst & Young (2004) state that as businesses move toward increasingly decentralized business models and with other external partnerships, it becomes even more difficult to retain control both in the management of shared resources and over the security of information. Furthermore, failure in PPPs can result from poor legal framework and enforcement, weak institutional capacity and PPP strategy, unrealistic cost and return on investment estimations, risk sharing arrangements are ineffective or inadequate.

There are also clear *ex ante* and *ex post* risks. The key *ex ante* risk is that in ensuring the right partners and that is there is no adverse selection of partners and the key *ex post* risk is that once partners are selected they are able to carry out the contracts effectively and are able to manage and subsume existing and emerging risks. The business behaviours, however, may vary between the partnership members (Vogel, 1996, Godard et al 2002) and the effectiveness of trust is critical (Seligman, 2000, Braynov and Sandholm, 2002) for partner robustness.

There are two faultlines in ensuring robustness: incentives for co-operation and the mantnance of the aggregation of experience. There are diverse views emphasising that the failure of properly incentivising. Even if the problem of incentives might be overcome by drastic legislation there are increasing questions as to the link between incentives and effectiveness.

There has been research done on the impacts of information sharing (Anderson, 2001, Gordon and Loeb, 2002, Gordon et al, 2003). There is a trade-off between the level of investment in information sharing and investing in partnerships and capital investment to prevent and mange risks of cyber network security. There is a positive correlation between information sharing and the number of partners and the level of equipment used – as the more the partners, there is more information and more diversification of information and more risks identified for management. There is however negative correlation between incentives for investment and information sharing where the firm size is smaller or the degree of concentration of firms in networks security is minor (Gal-or and Ghose , 2005) .

In markets with a low level competition, it can be assumed that each party brings with it a set of knowledge. However, this can be potentially dangerous, because it would potentially allow non robust collaborating parties a long life line resulting in ineffective partnerships. In an environment of fast-changing products and plenty of competitors, any partnership will need more than the reliance on its partners as there could be creeping self-interests (Allaz and Vila, 1993). It would need a system of usage greater than the aggregation of knowledge and experience of each of the partners to ensure that there is not only a joint value system but some form of governance mandating their involvement in maintaining network security.

5 How Can Governments Play a More Proactive Part in Managing Network Security?

Governments are expected to play a significant role where network security partnerships affect public services and are of public security. According to Cohen (1999, p.342, Ungoed-Thomas and Sheehan (1999).) governments are aware that the stakes and risks involved in the intelligence business are generally significant. There have been cases where hackers have targeted large international banks with large public databases, penetrating their security systems and holding these organisations to ransom. Public security is tantamount with governments needing to nip any network threats in the bud by involving companies to co-share risks and participate in the security management process. An interesting paradox has been identified by Mickhail (2008) between considerations of national security and the Socratic notion of misology (misplaced faith in those with unreliable substance). Partnerships exists where economic benefit becomes paramount over national security and Mickhail suggests can result in the "degeneration of rationality into subservience to commercial interests resulting in the rise of a fundamentalist brand of global capitalism that thrives on the corporatisation of national security, and which is giving rise to a new security world order[*sic*]".

Apart from government initiatives there are industry moves toward better governance of information and security including the formation of public private security bodies sucha the us National Cyber Security Partnership (NCSP). President Obama unveiled in May 2009 the Cyberspace Policy review more in managing civil liberties

Table 2. PPP risks in network security

Risks	PPP benefits	literature
Technological	Sharing of resources Knowledge and expertise and co-ordination of provision Risk transference and risk sharing.	de Langen (2004), Anderson , (2001), Gordon and Loeb (2002), Gordon et al (2003)
Operational	Sharing of resources Investment in capital spend scale, scope and diversity of project	de Langen (2004)
Partnership	Partnership controls	Vogel, (1996), Godard et al (2002) Cohen (1999), Ungoed-Thomas and Sheehan (1999). Allaz and Vila (1993)
Market	Co-operation arrangements with other companies in the market Incentivising	Handley-Schachler and Navare (2009) Chalfinch (2007) Gal-or and Ghose (2005), Allaz and Vila (1993)
National	Co-operation arrangements to manage security	INSA US and other governmental initiatives

than in monitoring private sector networks. Similarly other governments have been doing the same with Britain ahead of the game setting up the task force for "Regulating Cyberspace Better Regulation for e-commerce"; Finland setting up cyber defence centre among similar developments in other countries.

As part of the first level conceptual research there appear to be five critical risks arising where PPPs have certain benefits.

6 Conclusion

This paper emphasizes the increasing risk of cyberwar between the nations of the world and lack of preventive measures at various levels to alleviate the acceleration of tension due to such malicious acts. The role of Public Private Partnerships is highlighted as a measure to control such aggression. The resulting long-term, trust-based partnerships reduce coordination costs and information asymmetries, thus making strategic advantages possible (Stölzle and Heusler, 2003, pp. 174- 177).

Network security issues affect risk and cost sharing, uncertainty reduction; capacity for financing, access to complementary resources and skills and a more effective way to deploy resources. To enable this more controlled risk management and governance is required.

References

1. http://www.symantec.com/content/en/us/about/presskits/SES_report_Feb2010.pdf (last accessed on 12/04/2010)
2. Allaz, B., Vila, J.-L.: Cournot Competition, Forward Markets and Effciency. Journal of Economic Theory 59, 1–16 (1993)
3. Anderson, R.: Why information security is hard: An economic perspective. In: Proc. 17th Annual Comput. Security Appl. Conf. (December 2001)
4. Braynov, S., Sandholm, T.: Incentive compatible mechanism for trust revelation. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems, Bologna, Italy (July 2002)
5. Cohen, E.: The CIA and the declassification of history. International Journal of Intelligence and Counterintelligence 12, 338–345 (1999)
6. Gal-Or, Ghose: The Economic Incentives for Sharing Security Information. Information Systems Research 16(2), 186–208 (2005)
7. Godard, O., Henry, C., Lagadec, P., Michel-Kerjan, E.: Traité des nouveaux risques: Précaution, Crise, Assurance, Gallimard, Folio Actuel, Paris (2002)
8. Gordon, L.A., Loeb, M.: The economics of information security investment. ACM Trans. Inform. System Security 5(4), 438–457 (2002)
9. Gordon, L.A., Loeb, M., Lucyshyn, W.: Sharing information on computer systems security: An economic analysis. J. Accounting Public Policy 22(6), 461–485 (2003)
10. Mickhail, G.: National Security and the Misology-Misanthropy Paradox of Technology. Australia and the New Technologies: Evidence Based Policy in Public Administration, July 22, pp. 80–86. University of Wollongong Press, Wollongong (2008)
11. Rios, M.J., de Magalgaes, S.T., Santos, L., Jahankhani, H.: The Georgia's Cyberwar. In: Proceedings of the 5th International Conference on Global Security, Safety, and Sustainability (ICGS3), pp. 35–42 (2009)

12. Rios, J.M., de Magalgaes, S.T., Santos, L., Jahankhani, H.: The People's Republic of China – The Emerging Cyberpower. In: Proceedings of the 5th International Conference on Global Security, Safety, and Sustainability (ICGS3), pp. 138–144 (2009)
13. INSA (2009)
14. Handley Schachler, M., Navare, J.: Port risk management and Public Private Partnerships: factors relating to risk allocation and risk sustainability. World Review of Intermodal Transportation Research 3(1-2) (2010)
15. Seligman, A.: The Problem of Trust. Princeton University Press, Princeton (2000)
16. Ungoed-Thomas, J., Sheehan, M.: Hackershold city banks to ransom. The Sunday Times, p. 30 (September 19, 1999)
17. Vogel, D.J.: The study of business and politics. California Management Review 38, 146–165 (1996)

The Effect of Non-technical Factors in B2C E-Commerce
(A Case Study in Iran)

Ali Sanayei[1] and Reza Shafe'ei[2]

[1] Director of ITM Research Group, University of Isfahan, Iran
www.DrSanayei.com
[2] Assistant professor, University of Kurdistan

Abstract. As e-commerce grows across industries worldwide , business are building web sites for presence as well as for online business. It is more than transferring current business operations to a new medium. This situation requires explaining main models, changing infrastructures, and notice to customer needs as their vital rights. Whilst increasing numbers of firms have launched themselves on the Internet, they are trying to consideration of the strategic implications of developing, implementing or running a Web site. Global competition, laws, and customer preferences are among the issues being affected by e-commerce. In this study many factors that effect on e-commerce are considered these factors have no technical issue in nature. Companies related factors, customers' knowledge, customers' trust and customers' behavior are the main effective factors in development of B2C e-commerce. In this research we surveyed the mentioned aspects by offering questionnaire to experts of e-commerce for companies. The results show there is a meaningful relationship between perception, knowledge, trust and attitude of customers and the company's capabilities in the other side with B2C e-commerce development.

Keywords: Customers' knowledge, Customers' attitude and behavior, Customers' trust, Nontechnical factors, Companies related factors.

1 Introduction

The internet is the fastest-growing, most user-friendly, and most commercially popular technology to date. Anyone with a PC connected to the internet, a browser, and plug-in can surf the internet and download text, graphics, and even voice. Studies illustrated that around percent of Cisco's 800,000 monthly customer queries are handled via the Web. In this way customer orders are routed directly to contract manufactures, speeding product delivery and minimizing inventory.(Dave , 1999.177) The "buy and sell" aspect of internet commerce has attracted more media attention than any other networked activity to date.(Awad .2002, 43) Many companies provide configurable products on Internet to satisfy customers' diversified requirements. Most of business-to-consumer(B2C)e-commerce software systems use tree- or wizard-like approaches to guide customers in configuring a customized product on Internet web pages. However, customers may feel confused while they are selecting components of a product from option lists, since they are usually not familiar with the technical details of these components. (Luo et al .2008 ,

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 178–192, 2010.

1) Therefore web site evaluations (of content, organization, and technology) were posited as affecting the success (involving satisfaction, commitment, and trust) of a multi-dimensional web site. Pleasure was introduced as a key variable, mediating the relationship between web site evaluation and web site success. Many studies found that pleasure partially mediated the evaluations–success relationship and also found significant support for direct relationships between web site evaluations and success. Additionally, satisfaction was found to be instrumental in helping predict user commitment and trust that online shoppers placed on the site owner.(De Wulf . 2005 ,434) For surveying important of this aspects, we studied many factors can affect on e-purchasing via the web by customers. A favorable e-tail store image positively influences e-patronage intentions, which thus leads to e-loyalty.(Yun and Good,2007 . 4)

In this paper business to customer electronic commerce (B2C e-commerce) is examined in the context of the relationship between firms and their customers and the implications for organizational accountability. The technology of e-commerce determines what can be offered to customers, but only customers determine which of those technologies will be accepted. The authors argue that providing the highest customer delivered value by e-commerce can be viewed as making a real contribution to customers, through the Internet will be accepted by customers. Also there are some factors influence on business to customer deals that they have a nontechnical aspects in e-commerce processes. They are many subjects related to the companies, the extent of awareness, trust and attitude of e-customers.

1.1 B2C E–Commerce

Business-to-Consumer (B2C) e-commerce focuses on direct businesses between companies and final consumers (Dedhia, 2001; Lawrence, et al., 2000; Riggins and Rhee, 1998; Schneider and Perry, 2000; Ah-Wong, et al., 2001), it has a meaning that, the trading and transactional connection between an organizations website and an user (Dedhia, 2001; Lawrence, et al., 2000; Riggins and Rhee, 1998). Consumers are able to purchase goods and services such as books, computer products, music, at one time that is suitable to the purchaser. One of the key benefits of e-commerce is convenience, that is, day and night trading, 365 days of the year (Dedhia, 2001; Chen, Ingraham and Jenkins, 2001; Lohse and Spiller, 1998). Despite this benefit of e-commerce, would-be consumers are still concerned about purchasing over the Internet. Whilst ecommerce increases, so too, do trust concerns (Gray and Debreceny, 1998; Cheung and Lee, 2001; Urban, Sultan, and Qualls, 2000; Ernst and Young, 1999). In these studies, the main questions asked by would-be consumers were: Is this company real? Is this a trustworthy company? If I send credit card or bank information, is it safe? If I place an order, will I receive what I asked for? Will any problems I have be resolved quickly? There is a common theme to these questions, with a majority of them referring to the trust and risk in the trading relationship between the website and the consumer. Quantitative studies, on the Internet (UMR Insight, Ltd, 1999; Ministry of Economic Development, 2001) completed in New Zealand found that 11% of those surveyed had made a purchase over the Internet using a credit card (Ministry of Economic Development, 2001). The frequency of purchasing via the Internet was higher in Wellington (eighteen percent) amongst the ‘white-collar’ occupational group (seventeen percent) and those in the top income group (eighteen percent). Twelve percent

of males surveyed were more likely to have made a purchase than females (four percent). In 2000 twenty-one percent of those surveyed felt comfortable when shopping over the Internet and twenty one percent felt uncomfortable using a credit card (Ministry of Economic Development, 2001). were also identified as a group that felt distinctly uncomfortable giving their personal details via the Internet. Therefore, to encourage their presence on the Internet, there is a need to understand the factors that influence their trust interpretation of for online shopping.

With the advance of Internet technology, e-businesses are striving to reach an unprecedented large population and start to take on new forms. An e-business can be built on top of a wide range of e-business models [Singh 2002]. This paper selects business-to-consumer (B2C) Internet retailers and examines customers' attitude and patronage behavior toward them. Understanding what motivates customers to adopt and patronage Internet retailers is important because it is the key to Internet retailers' survival in the intensely competitive market. The competition comes not only from the e-commerce market, but also from alternative channels such as traditional retailers [Chen et al. 2002]. Compared with a traditional retailing environment, enticing customers to an Internet retailing environment is far more challenging. It is because that the online environment requires customers to make substantive behavioral changes in adopting and trusting e-commerce technologies and making informed decisions using technologies [Bhattacherjee 2000]. Therefore, obtaining knowledge about customers' attitude toward Internet retailers can help businesses develop effective marketing strategies for attracting and retaining customers and gain competitive advantage. There is a growing body of literature on examining customers' acceptance of Internet shopping [Meyer and Johnson 1995; Liang and Huang 1998; Devaraj et al. 2002; Kwak et al. 2002; Shim et al. 2002; Pavlou 2003]. Several factors, including customers' perception of convenience, product offerings, production information, site design, financial security of Internet stores, and trust, are found influential to customers' satisfaction with Internet shopping [Liang and Huang 1998; Szymanski and Hise 2000; Gefen et al. 2003; Vatanasombut et al. 2004]. Undoubtedly, improving the performance of those factors can potentially attract more customers and increase the effectiveness of Internet retailing services. In reality, however, due to resource constraints such as technical feasibility, cost feasibility, and organizational feasibility, it is almost impossible for Internet retailers to improve all of the factors affecting customers' perception simultaneously. This reveals a need to focus on a subset of key attributes that have been identified in the relevant literature.

1.1.1 Customer's Knowledge

The extent of visiting the web sites by customers has positive relationship to their knowledge and educations. Some studies had shown customers who can search in the web and look for their desires and find them in the world wide web, have a good ability to use it.

There is growing intent of those customers' acceptance of Internet shopping evidences. [Meyer and Johnson 1995; Liang and Huang 1998; Devaraj et al. 2002; Kwak et al. 2002; Shim et al. 2002; Pavlou 2003]. In the digital market, attracting sufficient online traffic in a business to customer Web site is vital to an online business's success. The changing patterns of Internet surfer access to e-commerce sites pose challenges for the Internet marketing teams of online companies. For e-business to grow,

a system must be devised to provide customers' preferred traversal patterns from product awareness and exploration to purchase commitment. Such knowledge can be discovered by synthesizing a large volume of Web access data through information compression to produce a view of the frequent access patterns of e-customers. This paper develops constructs for measuring the online movement of e-customers, and uses a mental cognitive model to identify the four important dimensions of e-customer behavior, abstract their behavioral changes by developing a three-phase e-customer behavioral graph, and tests the instrument via a prototype that uses an online analytical mining (OLAM) methodology. The knowledge discovered is expected to foster the development of a marketing plan for B2C Web sites. A prototype with an empirical Web server log file is used to verify the feasibility of the methodology. (Kwan et al, 2005) Rowley (2002) argues that customer knowledge is an important asset for all businesses. The rhetoric of e-business emphasizes the opportunities for knowing customers in the digital economy. This article sets the context with a brief summary of the key characteristics of the knowledge management paradigm. This is used as a platform for the formulation of the questions that form the core of this article: What customer knowledge do businesses require? What customer data can be collected? What are the challenges for translating data into information and knowledge? Can knowledge cultures be created in online customer communities? Whose knowledge is it anyway? How can knowledge assets be identified and managed in virtual organizations? How can customer knowledge from e-business be integrated with customer knowledge from other channels? Who needs customer knowledge anyway?

1.1.2 Customer's Trust

The concept of trust has been surveyed under various contexts (Cheung and Lee, 2001; Stewart, 1999; Choudhuri and Holbrook, 2001; Steinauer, Wakid and Rasberry, 1997; Hoffman, Novak and Peralta, 1999) and include trust in bargaining (Schurr and Ozanne, 1985), distribution channels (Dwyer, Schurr and Oh, 1987), industrial buyer-seller relationships (Doney and Cannon, 1997), partner cooperation in strategic alliances (Das, 1998) and the use of market research (Moore, Deshpande and Zaltman, 1993).

According to Lewicki and Bunker (1995), three theoretical perspectives exist. The first is the view of personality theorists, who conceptualize trust as a belief, expectancy or feeling that is deeply rooted in the personality of the individual. The second is the view of sociologists and economists, who see trust as a phenomenon within and between organizations and as the trust individuals put into those organizations. The third view is that of social psychologists, who characterize trust in terms of expectation and willingness of the trusting party to engage in a transaction (Lewicki and Bunker, 1995). For the purposes of this study, the definition offered by McKnight, Cummings and Chervany (1998: 459) has been adopted: “an individual’s beliefs about the extent to which a target is likely to behave in a way that is benevolent, competent, honest, or predictable in a situation.” The consumer needs to believe that the Web merchant is trustworthy before they purchase online. A consumer’s willingness to buy from an Internet seller is dependent on the their attitude towards the store, which, is affected by the seller’s ability to create consumer trust. Jarvenpaa, et al. (1999) proposed a model for the consequences of trust in an Internet store, seen in

Figure 1 below. According to Jarvenpaa, et al. (1999), the model suggests, that consumers' evaluation of stores' reputation and size affect their trust in a store. Higher trust will directly improve attitudes towards a store (Jarvenpaa, et al., 1999). Besides helping to shape attitudes, perceived risk might also have an independent, direct influence on the willingness to buy. According to the model, a consumer may be willing to buy from an Internet store that is perceived as low risk, even if the consumer's attitudes towards that merchant are not highly positive. Conversely, a consumer may not be willing to buy from a merchant perceived as being high risk, even in the presence of positive attitudes towards that merchant.

1.1.3 Customer Behavior and Attitude

If a company manages to improve its e-commerce transaction results, this will have a favorable effect on overall perceived quality and consumer attitudes. Usually customer behavior cannot be easily predicted , and it should be study very carefully with known tools which use increasingly in psychological analysis. In other words, the result of analyzing is as important as the process for a successful purchase by customer. It can be named as a behavior.(Alzola and Robaina.2007, 284) The effect of customer attitude on his or her purchase is deferent because of the difference in traditional business models, conventional consumer behaviors, and consumer expectations between countries and this feature can also a new complexity factor of e-commerce.(Wong et al . 2004, 68) Style of managers can change customer attitude in purchasing process when he or she will buy via the web. (Cope and Waddle .2001, 523). This situation occurs while the buyer search the goods or services to order. The ways of introducing of goods and services , showing of warranty methods , framework of web site, usability of site maps and other factors are structured by staff and they are affected by their bosses style. Electronic commerce (e-commerce) is examined in the context of the relationship between firms and their customers and the implications for organizational accountability. The technology of e-commerce determines what can be offered to customers, but only customers determine which of those technologies will be accepted. The author argues that providing the highest customer delivered value by e-commerce can be viewed as making a real contribution to customers, i.e. shopping through the Internet will be accepted by customers. Customer satisfaction is of critical importance when measuring perceived customer delivered value that is offered by e-commerce. Three main scales which play a significant role in influencing customer satisfaction are customer need, customer value and customer cost.(Lin .2003:202) The findings of a study suggest that operations-based competitiveness in e-commerce requires the development of a series of distinctive competencies, customer oriented program, customer attitude evaluation and prediction of customer behavior that those competencies are often related and mutually supportive, and that there is usually a linkage between distinctive competencies in e-commerce operations and the business strategy. (Da Silveira . 2003 : 200) Wen et al (2001 : 5) In their study mentioned that the rapid adoption of the Web as a commercial medium has caused firms to experiment with innovative ways of doing business. Those firms that effectively market themselves on the Web have a distinct advantage. Also they studied the importance of customer consideration in the companies planning as a vital factor to achieve the goals in progress. Customer attitude is a characteristic that if it will known, company can provide goods and services to him appropriately. Attitude

has three important component; the cognitive , the effective and the behavioral component. This can be learned or changed after the customers gather information from some sources they will.(Loudon and Bitta, 2004 :425) Evidence or habit rather than attitude supporting that past behavior frequently predict future behavior has since then accumulated(e.g., Bagozzi, 1981; Bentler & Speckart, 1979; Fredricks & Dossett, 1983; Kahle, 1984;Kahle et al., 1981; Landis et al., 1978; Mittal, 1988; Wittenbraker et al., 1983). Yet, it is argued that habitual behavior sequences are perhaps frequently functional for obtaining certain goals or end states (Bargh & Gollwitzer,1994).

1.1.4 Non Technical Factors

There are many variables can affect on buyers in the web such as organizational factors, managerial skills and so on. The following researches indicated the other effective factors in electronic dealing. De Wulf et al (2006) developed and empirically validated a process model of web site success in an online shopping context by identifying the role of pleasure as a key mediating variable. They believed that the web site evaluations (of content, organization, and technology) were posited as affecting the success (involving satisfaction, commitment, and trust) of a multi-dimensional web site. Pleasure was introduced as a key variable, mediating the relationship between web site evaluation and web site success. The pleasure partially mediated the evaluations–success relationship and also found significant support for direct relationships between web site evaluations and success. Additionally, satisfaction was found to be instrumental in helping predict user commitment and trust that online shoppers placed on the site owner. The old notions of management are totally ineffective and a new style, focused on "leadership", is required – but what style of leadership? The findings of an audit on leadership styles indicate that organizations that have successfully integrated e-commerce have exhibited a unique approach Cope and Waddell (2001). Using a change management matrix, which determines the impact of change versus leadership style within an organization, 182 Australian managers were audited and positioned within this matrix. It was found that within the most successful organizations, leaders had a distinctive style that facilitated the appropriate change and established a conducive e-commerce environment. Also poor planning practice is one of the reasons for the inferior results of some e-commerce ventures. In this paper, it is suggested that a strategy-based e-commerce planning model containing seven specific dimensions, with strategy at its core, should be considered when entering into a new e-commerce venture. Implications for managerial practice upon the adoption of such a planning model are also discussed.

1.1.5 Conceptual Modeling

In our study, a model is designed to examine all of variables and illustration of their correlations. In this model we illustrate the relation between main variables. (Figure 1) It was based on many important researches.

1.2 Hypotheses

The above proposed model deals with many factors that have positive effect on B2C e-commerce development. Those factors are known as non-technical aspects. For analyzing of the related variables our focus is on the behavior of buyer. The model preserves the original background of other studies. In the context of notice from the

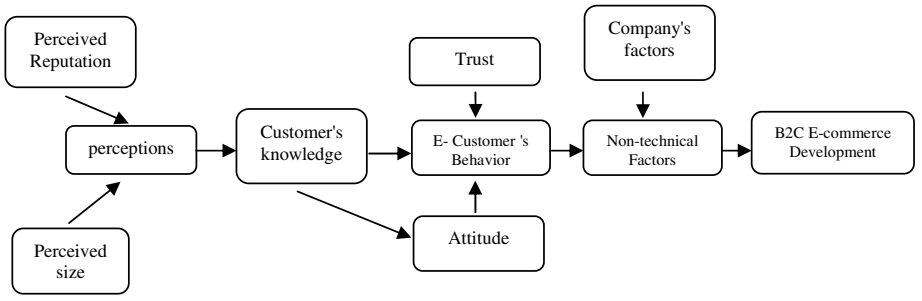


Fig. 1. Model of analyzing of role of the non-technical factors in B2C e-commerce development

firms to e-commerce according to the customer's view as their attitude functions, we should test the model for surveying many aspect of our assumptions. Accordingly, we have the following hypotheses:

H_1: Customer's knowledge have positively affected on E-customer's behavior.
H_2: Customer's trust and attitude have positively affected on E-customer's behavior.
H_3:E-customer's behavior as a Non-technical factor has positively effect in B2C e-commerce.
H_4:The company performance as a Non-technical factor has positively effect in B2C e-commerce.
H_5: Each of the mentioned factors has different effect in B2C e-commerce development.

2 Methodology

A descriptive survey methodology was used. A questionnaire was prepared to measure the proposed model, with items that they were derived of the prior studies . At the first step the responders specified the extent of effects of each factor in the midrates variables. The questionnaire was pretested by faculty members. The questionnaire was offered to the experts of e-commerce in the universities. The demographics' features of them are shown in Table 1; and other analyzing presented in the following tables .

Table 1. Sample Demographics

Total characteristics	
Gender: Female	40%
Gender: Male	60%
Age Mean	34 years
Standard deviation	2.2
Years of experiences Average	8 years
Organizational position average	Expert (faculty members)

3 Analysis and Results

Confirmatory factor analysis was used to test the measurement model and establish convergent and Lisrel software was used for testing the relation between variables and factors based on SEM method . The reliabilities of all constructs exceeded the minimum acceptable Cronbach's alpha level of 0.878, indicating internal consistency. The most of the scales had Cronbach's alpha greater than 0.90. The first analytical approach was factor analysis. In that method we wanted to confirm four considered factors which have identified as the customer's factors and one factor for organizations.

For confirmation of the first hypothesis it was necessary to show validity of the considering factors.

All operation of the mentioned process were shown in Table 2.

Table 2. The results and analyzing of the main factors

Main Factors	Component Factors	Question Numbers	Extraction Average	Total Variance Explained	Eigen Value
Customer's Factors	Perception	Q1-Q10	.771	23.235	1.677
	Knowledge	Q11-Q20	.643	11.121	1.781
	Trust	Q21-Q25	.878	4.684	1.917
	Attitude	Q26-Q30	.712	4.231	1.410
Companies' Factors	Cultural	Q29-Q50	.756	4.094	1.637
KM=.812	Bartlett's Test of Sphericity:7656.9 Sig=.000				

The variables of companies and customer's factors that were the main assessable factors in this research are evaluated. They are surveyed by five factors that shown in table 2, our method was confirmative factor analysis and the results explained an agreeable consequence. The extraction averages were near to one, the total variance

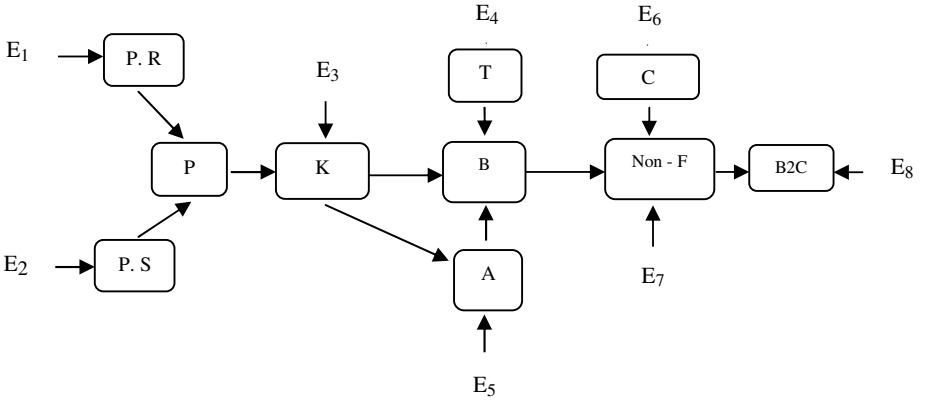


Fig. 2. Showing the Conceptual Model of the research through SEM model diagram

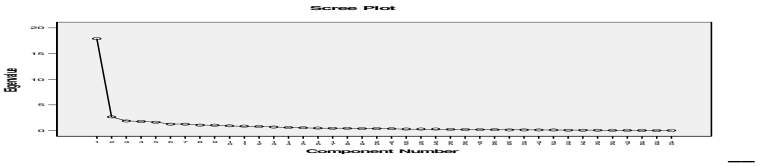


Fig. 3. The Scree plot of confirmative factor analysis

explained were further than .50 and the Eigen values were more than one. Also the results illustrated the KM score and the Bartlett's test were significant. Consequently according to the table and figure 2 the mentioned factors are appropriate to study of the organizational structure components.

The analyzing of the hypotheses

H1: Customer's knowledge have positively affected on E-customer's behavior.

Table 3. Analyzing of the customer's factors

ANOVA

The factors		Sum of Squares	Mean Square	F	Sig.
Perception	Between Groups Within Groups Total	56.239 12.547 98.856	1.138 .031	32.8	.032
Knowledge	Between Groups Within Groups Total	44.161 24.005 79.243	1.562 .083	28.25	.021

In the first step we assessed the effect of two variables from customers, its perception and knowledge about using e-business. According to table 3, there is a significant relationship between customer's perception and knowledge in use of its behavior.

Table 4. Analyzing of the descriptive situation

Descriptive Statistics

The factors	Minimum	Maximum	Mean	Std. Deviation
Perception	1.64	4.10	3.223	.60
Knowledge	1.76	4.21	3.045	.61

According to the table 4, the extent of the studied expert's opinion about the effect of perception and knowledge in customer's behavior is further than average level. Then we can conclude, H_1 is confirmed.

H_2: Customer's trust and attitude have positively affected on E-customer's behavior.

Table 5. Analyzing of the effect of trust and attitude variables on the related factor

Model Summary(a)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	F	Sig.
1	.91(a)	.828	.828	.1324	216.113	0.000

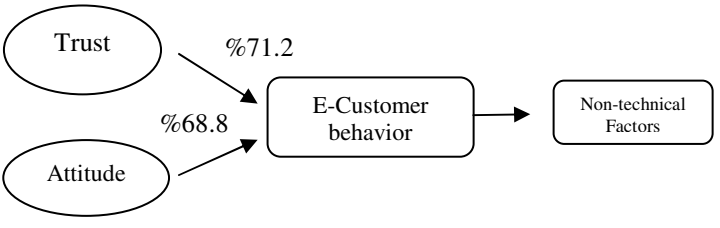


Fig. 4. The path analysis between the human-relate factors and organizational structure

The above results show that there are significant relationship between trust and attitude as the other individual factors and e-customer behavior. In the operation it is seen that the amount of Durbin-Watson was 1.67 that it is more than 1.5 and F = 445.321 and sig.=0.000. Therefore H_2 is confirmed. In the other hand, the other customer's factors (trust and attitude) have positively affect on E-customer behavior.

H_3:E-customer's behavior as a Non-technical factor has positively effect in B2C e-commerce.

Table 6. Analyzing of importance of factors

Coefficients(a)

Model		Unstandardized Coefficients		t	Sig.
		B	Std. Error		
1	(Constant)	5.124	.134	21.226	.000
	Perception	1.213	.034	11.225	.000
	Knowledge	3.032	.048	13.429	.000
	Trust	2.234	.063	12.421	.000
	Attitude	1.214	.033	14.211	.000

a Dependent Variable : acceptance

According to the table 6, the extent of the important of risks in effect on acceptance of e-banking is different. In the other hand the table showed the performance risk has further effect in all of those risks. Other table illustrated a ranking of the important of risk's effect.(table 6)

Table 7. Ranking of the customer's variables

Ranks

Risks	Mean Rank	Ranking
Knowledge	5.14	1
Trust	4.32	2
Perception	3.22	3
Attitude	2.78	4
Chi- square:130.090	df:4	Sig:0.00

According to table 7 the maximum important of customer related factors is knowledge and the lowest degree is for attitude from view of the experts.

H_4:The company performance as a Non-technical factor has positively effect in B2C e-commerce.

Table 8. Analyzing of the effect of company factor

Model Summary(a)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	F	Sig.
1	.91(a)	.828	.828	.1324	216.113	0.000

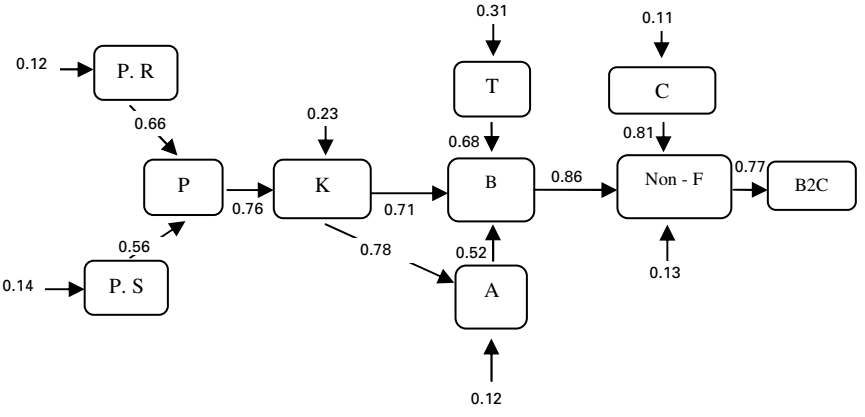
According to tables 8 the performance of organizations in e-business against their customers has positively effect in continues of this relation. In the other hand to a good B2C commerce via the web, the companies have main cause to attraction of e-customers in new communication.

H_5: Each of the mentioned factors has different effect in B2C e-commerce development.

For surveying the relation between the two main factors, customer's and company aspect, we surveyed the amount effect of main factors on important variable. For this reason we did this analysis by SEM method by Lisrel software. In the figure 3, the resulted model has shown. In that model we illustrated the relation among three group of variables, customer group and the organization variables.

The analytical model shows there are meaningful relationship between the main group variable as customer's and the organization's variable with B2C e-commerce. The above model assessed by Lisrel software which it use for analyzing SEM. Perception of customers about e-commerce 0.76, knowledge, 0.71, attitude, 0.52, trust,

0.68, and the company performance has 0.81 of the total effects in B2C e-commerce development in Iran, in general all of them have 0.77 of the whole effects on the e-commerce. The perceived reputation and the perceived size variables are evaluated, the result showed that they have a meaningful influence in the customer's perception about e-commerce.



Chi-Square=21.34, df=12, P-value=0.212, RMSEA=0.305

Fig. 5. The SEM model for analyzing the relation between the factors

Table 9. Measuring the structural model

No	Hypothesis path	Path coef-ficient	t-Value	p-Value	Supported
--	P.R → P	0.66	1.34	0.00	--
--	P.S → P	0.56	2.06	0.00	--
H_1	P → K	0.76	3.45	0.00	Yes
--	K → A	0.78	1.16	0.00	--
H_1	K → B	0.71	1.24	0.00	Yes
H_2	A → B	0.52	2.13	0.00	Yes
H_2	T → B	0.68	2.42	0.00	Yes
H_3	B → Non	0.86	3.56	0.00	Yes
H_4	C → Non	0.81	2.78	0.00	Yes
H_5	B&C → B2C	0.77	1.52	0.00	Yes

4 Conclusion

This paper aims to develop an extended model to predict and explain e-customers' behavior with regard to their individual features. The proposed model incorporated per five hypothesis and provide a more comprehensive investigation covering the

positive and negative aspects of factors that influence in B2C. The results show that the proposed model has good clarifying influence and confirms its strength in predicting customers' attitude of using. This study identified two aspects of perception(reputation and size of perception) that influencing on consumers' knowledge in online buying, it is important to recognize the cultural and regional limitations of these findings in one side and the national side in the other side. Moreover, according to other researches individual's understanding of trust differs between people and is likely to affect the perceptions of the presence of trust as well as the evaluation of the trust and attitude. In other words, the customers' acceptance of e-commerce may be indirectly influenced by personal features. However, this results needs further investigations and studies. Hence, the replication of this study on a wider scale with different national cultures is essential for the further generalization of the findings. Analysis further reveals that accessibility of internet, awareness of e-purchasing, and customers' intention to use new technology are the factors that significantly affected the usage of e-commerce in Iran. As we explained already e-commerce is defined as the new way to delivery of new and traditional products and services directly to customers through electronic, interactive communication channels trying to reach a widespread using of it, want to decrease any potential risks that imagine in the thinking of customers. For this reason website of companies should plan to provide this view that the e-commerce includes the systems that enable customers, individuals or businesses, to access goods, transact business, or obtain information on financial products and services through a public or private network, including the internet or other tools of servicing in new method. According to the results of this study customers who access e-commerce services using an intelligent electronic device, such as a personal computer , personal digital assistant , automated teller machine, have really understanding of that electronic financial devices. In the other hand, customers with repeatedly access to online facilities can accept very quickly e-commerce. That issues examiners should consider when reviewing informational websites include: potential access to confidential financial institution or customer information if the website is not properly isolated from the financial institution's internal network; potential liability for spreading viruses and other malicious code to computers communicating with the institution's website; and negative public perception if the institution's on-line services are disrupted or if its website is defaced or otherwise presents inappropriate or offensive material. The attitude of customers also can affect them when they want to access to the internet. According to our suggested model the company mangers should focus on some factors that influence people as very important factors in e-commerce. Other important factor that effect in e-commerce is the company's aspect. This factor includes website design, responsibility against customers, notice to the customer's views and provide what they want from B2C e-commerce.

References

1. Ajzen, I.: The theory of planned behavior. Organizational Behavior and Human Decision Processes 50, 179–211 (1991)
2. Azar, H.: The effect of relative thinking on firm strategy and market outcomes: A location differentiation model with endogenous transportation costs. Journal of Economic Psychology 29, 684–697 (2008)

3. Bargh, J.A., Gollwitzer, P.M.: Environmental control of goal-directed action: Automatic and strategic contingencies between situations and behavior. In: Nebraska Symposium on Motivation, vol. 41, pp. 71–124 (1994)
4. Bagozzi, R.P.: Attitudes, intentions and behavior: A test of some key hypotheses. Journal of Personality and Social Psychology 41, 607–627 (1981)
5. Bentler, P.M., Speckart, G.: Models of attitude-behavior relations. Psychological Review 86, 452–464 (1979)
6. Beatty, S.E., Coleman, J.E., Reynolds, K.E., Lee, J.: Customer sales associate retail relationships. Journal of Retailing 72, 223–247 (1996)
7. Bendapudi, N., Berry, L.L.: Customers' motivations for maintaining relationships with service providers. Journal of Retailing 73, 15–37 (1997)
8. Bennett, R.: Relationship formation and governance in consumer markets: transactional versus the behaviorist approach. Journal of Marketing Management 12, 417–436 (1996)
9. Berry, L.L.: Relationship marketing of services—growing interest, emerging perspectives. Journal of the Academy of Marketing Science 23, 236–245 (1995)
10. Berry, L.L., Gresham, L.G.: Relationship retailing: transforming customers into clients. Business Horizons 29, 43–47 (1986)
11. Bitner, M.J.: Building service relationships: it's all about promises. Journal of the Academy of Marketing Science 23, 246–251 (1995)
12. Rungie, C., Laurent, G., Dall'Olmo, F., Morrison, D., Roy, T.: Measuring and modeling the (limited) reliability of free choice attitude questions. Intern. J. of Research in Marketing 22, 309–318 (2005)
13. Chulmin, K., Sounghie, K., Subin, I., Changhoon, S.: The effect of attitude and perception on consumer complaint intentions. Journal of Consumer Marketing 20, 352–371 (2003)
14. De Wulf, K., Odekerken-Schr, G.: Assessing the impact of a retailer's relationship efforts on consumers' attitudes and behavior. Journal of Retailing and Consumer Services 10, 95–108 (2003)
15. Donio, J., Massari, P., Passiante, G.: Customer satisfaction and loyalty in a digital environment: an empirical test. Journal of Consumer Marketing 23, 445–457 (2006)
16. Fishbein, M., Ajzen, I.: Belief, attitudes, intention, and behavior: An introduction to theory and research. Addison-Wesley, Reading (1985)
17. Fredricks, A.J., Dossett, D.L.: Attitude-behavior relations: A comparison of the Fishbein-Ajzen and the Bentler-Speckart models. Journal of Personality and Social Psychology 45, 501–512 (1983)
18. Ghosh, A.: Retail Management. The Dryden Press, Forth Worth (1994)
19. Hernandez, B., Jimenez, J., DeHoyos, J.: Differences between potential, new and experienced e-customers: Analysis of e-purchasing behavior. Journal of Internet Research 18, 248–265 (2008)
20. Kahle, L.R.: Attitudes and social adaptation: A person-situation interaction approach. Pergamon, Oxford (1984)
21. Kahle, L.R., Klingel., D., Kulka, R.A.: A longitudinal study of adolescent attitude-behavior consistency. Public Opinion Quarterly 45, 402–414 (1981)
22. Kahneman, D.: Maps of bounded rationality: Psychology for behavioral economics. American Economic Review 93, 1449–1475 (2003)
23. Konrad, J., Theerasak, T.: Exploring Trust in B2C E-Commerce an Exploratory Study of Maori Culture. In: NEW ZEALAND ECIS 2002, Gdańsk, Poland, June 6-8 (2002)
24. Lior, F., Aviv, Z., Dov, T.: The effectiveness of online customer relations tools: Comparing the perspectives of organizations and customers. Journal of Internet Research 18, 211–228 (2008)

25. Lai, C.- H.: The research on the relationship among product attribute of car industry, brand image and brand loyalty take Toyota Taiwan as an example, MA, grand management, China (2006)
26. Loudon, D., Bitta, A.: Consumer Behavior. Mc Graw-Hill, New Delhi (2004)
27. Lee, J., Do-Hyung, P., Han, I.: The effect of negative online consumer reviews on product attitude: An information processing view. Journal of Electronic Commerce Research and Applications 7, 341–352 (2008)
28. Landis, D., Triandis, H.C., Adamopoulos, J.: Habit and behavioral intentions as predictors of social behavior. The Journal of Social Psychology 106, 227–237 (1978)
29. Morgan, R.M., Hunt, S.D.: The commitment–trust theory of relationship marketing. Journal of Marketing 58, 20–38 (1994)
30. Mittal, B.: Achieving higher seat belt usage: The role of habit in bridging the attitude-behavior gap. Journal of Applied Social Psychology 18, 993–1016 (1988)
31. Miller, R.L.: Weber and the consumer. Journal of Marketing 26, 57–61 (1962)
32. Malley, L., Tynan, C.: Relationship marketing in consumer markets: rhetoric or reality? European Journal of Marketing 34(7), 797–815 (2000)
33. Ole, B., Gärling, Fuji, S.: Empirical Tests of a Model of Automobile Choice Incorporating attitude, Habit, and Script. Paper presented at the Urban Transport Systems conference, Lund University, Sweden, June 7-8 (1999)
34. Palvia, P.: The Role Of Trust In E-Commerce Relational. Exchange: A Unified Model. Journal of Information & Management, S0378–7206 (2009) (accepted manuscript)
35. Parasuraman, A.: Reflections on gaining competitive advantage through customer value. Journal of the Academy of Marketing Science 25, 154–161 (1997)
36. Sanayei, A., Shafeai, R.: Vendors rating and its effect in E-customers with regard to supply chain management and E-Security. In: Proceeding of the 3rd Annual ICGeS, London (April 2006)
37. Sanayei, A., Shafeai, R.: The Use of Integrated Method of The Fishbein's Attitude Model and Customer-Oriented Requirements to Improve E-customer Behavior and Attitude (Case study: Automobile vending system in Iran). Journal of international marketing and research, 34 (2009)
38. Sanayei, A., Shafeai, R.: Analyzing the extent of notice to customer attitude in the organizational structure of vendors and its effect on e-buyer's trust (Case study in Iranian car producers). In: 5th International Conference on Global Security, Safety, and Sustainability, ICGS3, London, UK, 1-264-75 (September 2009)
39. Triandis, H.C.: Interpersonal behavior. Brooks/Cole, Monterey (1977)
40. Wittenbraker, J., Gibbs, B.L., Kahle, L.R.: Seat belt attitudes, habits, and behaviors: An adaptive amendment to the Fishbein model. Journal of Applied Social Psychology 13, 406–421 (1983)
41. Woodruff, R.B.: Customer value: the next source for competitive advantage. Journal of the Academy of Marketing Science 25, 139–153 (1997)

Self-monitoring Composite Rods for Sustainable Construction

Cristiana Gonilho-Pereira, Emilija Zdraveva, Raul Fangueiro, S. Lanceros-Mendez, Said Jalali, and Mário de Araújo

University of Minho, School of Engineering, Campus de Azurém,
4800-001 Guimarães, Portugal
cristiana.pereira@civil.uminho.pt, emilijia_zdraveva@net.hr,
rfang@det.uminho.pt, lanceros@fisica.uminho.pt,
said@civil.uminho.pt, maraujo@det.uminho.pt

Abstract. This paper presents the development and properties assessment of braided reinforced composite rods (BCR) able to both reinforce and monitor the stress state of concrete infrastructures. The research study aims at understanding the tensile behaviour and self-monitoring ability of composite rods reinforced by a textile structure – braided structure with core reinforcement – for civil engineering applications, namely for concrete internal reinforcement, as a steel substitute, in order to improve structures safety and sustainability. Seven types of braided composite rods have been produced using an author patented technique based on a modified conventional braiding machine. The tensile properties of the braided reinforced composite rods were evaluated in order to identify the type(s) of fibre(s) to be used as core reinforcement. BCR have been tested under bending while the variation of the electrical resistance was simultaneously monitored.

Keywords: sustainability, composite rod, tensile, self-monitoring, concrete.

1 Introduction

The concrete construction industry deals every day with the deterioration of concrete structures which compromises its security, safety and construction sustainability. Nowadays a large number of bridges, buildings and other structural elements require rehabilitation and repair and its maintenance have become an increasingly serious problem.

The corrosion of steel reinforcing rebar is the dominant cause of concrete structure degradation. The most effective way to prevent corrosion of steel rebar is the use of a corrosion resistant reinforcing material, such as fiber-reinforced-polymer (FRP) composites. The types of fiber-reinforced-polymer composites best suited for the reinforcement of concrete are those providing high strength, high stiffness, and environmental compatibility with concrete.

Nevertheless, the interest in the sustainability of concrete structures has increased and monitoring and maintaining their safety has become a main goal. To achieve this main goal monitoring systems that can be applied to the reinforced concrete elements

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 193–201, 2010.

are required. The damage sensing is conventionally performed by attached or embedded damage sensors, such as optical fibers, acoustic sensors, etc. however these sensors have limited application because of high cost, low durability, and limited sensing volume and spatial resolution. One solution is that the materials themselves can possess a self-diagnosing function for fracture; thus, strong and heavy design, complex and expensive equipment and numerous sensors becomes unnecessary, the so called self-diagnosing structural materials [1, 2].

Structural materials have evolved from materials that are mechanically strong (such as steel) to materials that are both strong and lightweight (such as composite materials) and most recently to materials that are both strong and self-monitoring [1]. By definition, a self-monitoring material is one which can sense its own strain and damage. It can be considered a smart material. However, in contrast to smart materials such as optical fibers, piezoelectric sensors, etc., the self-monitoring materials are themselves structural materials. Thus, instead of structures rendered smart by embedded or attached sensors, self-diagnosing structural materials are intrinsically smart, so there is no need of embedded or attached sensors. For example, the basic principle of the carbonaceous smart structural material to detect strain or damage lies in the electrical conductivity of the carbon fibers, as already known from the literature [2]. As the carbon fibres are electrically conductive, the composite itself can exhibit electrical properties, which will depend upon strain, damage and temperature. The self monitoring material will, in this way, provide determination of the strain or damage by measuring the change in the electrical resistance during real time loading [3].

Most commercial FRPs are rod-like elements that are pultruded, shaped, and treated so that surface texture provide mechanical adherence with concrete [4]. Besides pultrusion, fibre reinforced composite rods can also be produced using braiding techniques [5]. Braiding is a low cost technique allowing in-plane multiaxial orientation, conformability, excellent damage tolerance and allows core reinforcement. Moreover, braiding allows the production of ribbed structures and a wide range of mechanical properties may be improved when the core braided structures are reinforced with the appropriate type of fibers [6].

The current work is concerning the development of braided reinforced composite rods for civil engineering applications, namely for concrete internal reinforcement and monitoring.

2 Experimental Work

The objective of the experimental work presented in this paper, is the evaluation of the influence of the type of core reinforcement fibres on the tensile behaviour of braided composite rods and the assessment of its monitoring capabilities.

2.1 Braided Composite Rod Production

Braiding technique is one of the most ancient production processes of textile structures. The basic principle of braiding is the mutual intertwining of yarns. Braids are fibrous structures resulting of the yarns crossing in diagonal direction and can be tubular or flat, namely if they have round/oval cross section or not.

Core reinforced braided structures are braided tubular structures presenting, beside two systems of yarns moving helically, a third one that introduces yarns on the braid axial direction. This third system of yarns may be composed by different types of fibres, namely natural or man-made. The axial reinforcement fibres are responsible for the mechanical performance of core reinforced braided structures. The influence of the braided structure itself is rather poor.

Braided composites rods are produced in a conventional braiding machine with minor modifications, developed by the Fibrous materials Research Group, at University of Minho, allowing its impregnation in a polymeric matrix. Hence, using braiding technology with minor adaptations, ribbed composite rods may be produced in a single step according to an author patented technique based on a modified conventional braiding machine.

2.2 Raw Materials

Seven different braided composite rods were produced using polyester fibres for the braided structure production, E-glass, carbon and HT polyethylene fibres as braided structure core reinforcement, and a polyester resin was used for the core reinforced braided structure impregnation. Braided composite rods were produced maintaining the braided structure geometry and linear density and varying the type of core reinforcement fibre, according to Table 1.

Braided composite rods were reinforced with a single type of reinforcement fibres as well as with two and three types of fibres, varying the percentage of each one. Table 1 presents the percentage of each type of fibre used as core reinforcement over the total linear density of the core reinforcement.

Table 1. Braided composite rods composition

Rod type	Type of core reinforcement fibre		
	E-Glass fibre [%]	Carbon fibre [%]	HT polyethylene fibre [%]
1	100	-	-
2	77	23	-
3	53	47	-
4	-	100	-
5	50	45	5
6	52	45	3
7	75	22	3

The type and the amount of fibre were chosen to compare the tensile behavior of composite rods, as showed in Table 1. Rod type 2, 3 and 4 were tested regarding their monitoring ability based on the electrical resistance measurement during simultaneous application of a deformation in a cyclic three-point bending test.

2.3 Properties Evaluation

Table 2 presents the rod diameter and the fibre mass fraction of the reinforcement fibres of each rod produced. In order to evaluate the mass fraction of the different braided composite rods produced, tests were conducted according to the Portuguese Standard NP 2216/1988 (determination of mass loss by calcinations of glass fibre reinforced plastics).

Rods diameter varies from 5.27 to 6.40mm and the mass fraction of the core reinforcement fibres ranges from 31.8 and 40.6 %. According to Table 2, there is no relationship between the rod diameter variation and the volume fraction of the core reinforcement fibres. Therefore, the resin content varies from rod to rod.

During the curing period of the polyester resin, the core reinforcement fibres were subjected to a pre-load of 100N. In order to evaluate the mechanical performance of the different braided reinforced composite rods produced, tensile tests were carried out according to ASTM D 3916-94 standard, with a crosshead speed of 5 mm/min. A post-load of 50KN was applied to the rods prior to performing the tensile tests. Table 3 presents the average values of the tensile test results obtained for each rod type.

Table 2. Braided reinforced rods physical properties

Rod type	Rod diameter [mm]	Mass fraction (reinforcement fibres) [%]
1	5,50	40,6
2	5,27	35,3
3	5,75	31,8
4	6,40	33,3
5	6,00	35,6
6	5,98	32,7
7	5,78	33,7

Table 3. Tensile test results obtained for the different braided reinforced composite rods

Rod Type	Tensile strength [MPa]	Extension at failure	Tensile strength at 0.2% [MPa]	Modulus of Elasticity [GPa]
1	485,35	0,01701	110,73	55,36
2	766,70	0,01416	157,05	78,52
3	740,41	0,01178	148,96	74,48
4	747,77	0,01183	192,58	96,29
5	679,45	0,01105	167,84	83,92
6	652,77	0,01098	162,17	81,09
7	690,99	0,01438	146,40	73,20

The testing procedure carried on the BCR to evaluate the monitoring ability was based on the electrical resistance measurement during simultaneous application of a deformation in a cyclic three-point bending test. Cyclic three-point bending tests were carried on a Universal Testing Machine – Autograph IS (Shimadzu) 500N, with a crosshead speed of 0,3 mm/min. The electrical resistance measurement was carried on a digital multi-meter (Agilent, 84401A). The electrical signal was acquired through golden wires attached to the cross section of the samples with silver paint.

Representative examples of the two types of behavior obtained for the mechanical and electrical results from the cyclic loading three-point bending tests and the simultaneous electrical resistance measurements are presented in Figures 1 and 2 for the BCR samples tested. Figure 1 presents the increase in the electrical resistance with increasing displacement. Figure 2 shows curves showing the decrease of the electrical resistance with increasing deformation. In general, the electrical resistance during loading and unloading increases linearly at lower displacement values and nonlinearly at higher ones. In the case of the inverse response, during decrease of the electrical resistance and deformation increase, the nonlinearity is less evident. Whether it was a reverse or inverse response in both cases the tested sample showed the change of the electrical resistance in proper compliance with the change of its deformation. The main factors influencing the type of response of each sample was the relative position of the fibers and the resin inside the rods, or more precisely the carbon fiber placement along the length of the rods.

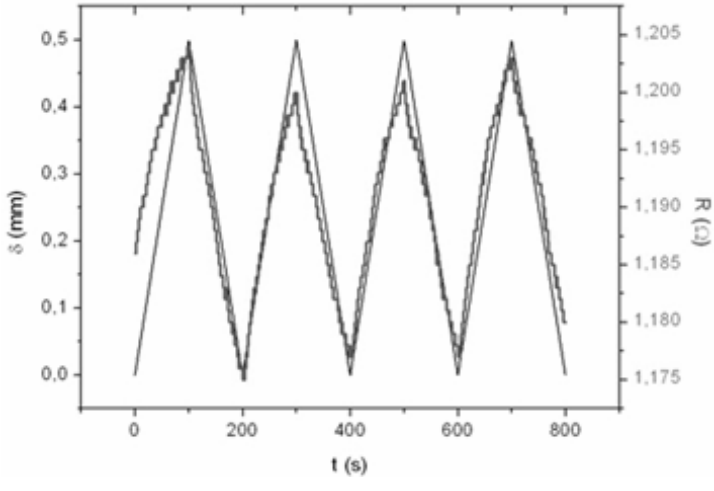


Fig. 1. Displacement, resistance change and time dependence for type BCR_4 (100% of carbon); Positive response of tested samples

The cross-sections present variations in the carbon fiber placement once the fibers are not placed uniformly on one side of the cross-section and, more important, the distribution is not the same on both cross-sections of a rod sample. In this situation, increasing of the resistance with increasing deformation occurred when the carbon fibers are placed on the tensile side and the opposite behavior when their placement is

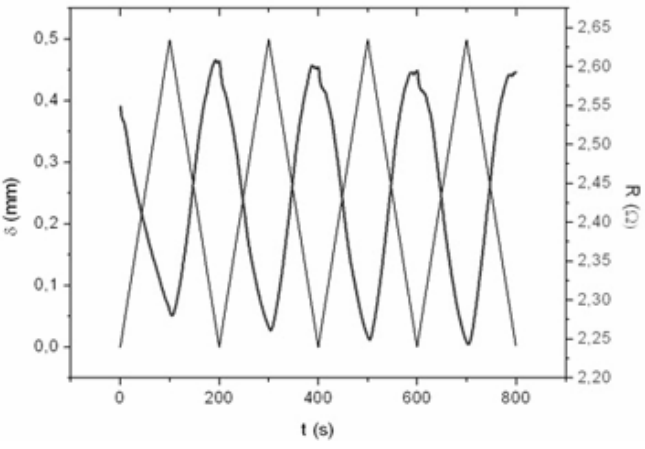


Fig. 2. Displacement, resistance change and time dependence for type BCR_2 (77% of glass, 23% of carbon); Negative response of tested samples

on the compression side of the bending rod. This issue indicates the relevance of controlling carbon placement uniformly along the length of the rod. This is a difficult step on the production process, which sometimes results in non-uniformity along the length of the rods as the braiding process rotates the fibers.

On the other hand, for each rod, the cycles are reproducible, confirming the reliable sensing property of the rods. The difference between the peak values in each of the cycles is around 0,01 Ω to 0,02 Ω.

The range of the initial electrical resistance for the BCR 2 varies from 1,46 Ω to 3,92 Ω, for BCR 3 from 1,11 Ω to 1,50 Ω and for BCR 4 from 0,80 Ω to 1,37 Ω.

Furthermore, for the comparison of the sensing behavior differentiation due to carbon fiber content, the strain ε ($\times 10^{-2}$) and the fractional resistance change $\Delta R/R_0$ of the three types of BCR are presented in Table 4. The strain is calculated from the displacement and the fractional resistance change is calculated from the electrical resistance change, the values from the two parameters are presented over time of 100, 300, 500 and 700 seconds.

As can be seen in Table 5, the gage factor for BCR 2 is almost five times higher than that for BCR 3 and BCR 4. The gage factor of type BCR 3 is slightly higher.

The gage factor, known as the strain-sensing factor, shows the sensing behavior of the composite rod samples. It increases with the decreasing of the carbon fiber percentage. This means that BCR 2 has the most reliable monitoring behavior.

It is interesting to compare the resulted gage factors of the BCRs with other materials for the same applications. For example in a study of a carbon nanotube strain sensors, it was investigated that the range of the gage factor was between 1 and 5, depending on the percentage of the single walled carbon nanotube (SWCNT) polymer composites, and better sensitivity was established in the range of 3 to 10wt% of the SWCNT in the polymer [6].

Another study, investigating a multi walled carbon nanotube (MWCNT) films used as strain sensing material, reported that the calculated gage factors were 2, 3,09 and 3,76, respectively, for 3 types of samples [7].

Taking these examples into comparison with the investigated samples in this study, it is evident that the gage factors calculated for the composite rods are all less than 1, much smaller than the example studies set forth above. The reason for these small values of the gage factors would be, as mentioned, the higher percentage of the carbon component. On the other hand, the main advantage of the present materials is the superior reinforcing capabilities and therefore the combination of reinforcing and sensing capabilities.

Table 4. Strain and fractional resistance change of BCR (mean values)

	Cycle No. t (s)	**1 100**		**2 300**		**3 500**		**4 700**	
Rod Type	Sample No.	$\varepsilon \times 10^{-2}$	$\Delta R/R_0$	$\varepsilon \times 10^{-2}$	$\Delta R/R_0$	$\varepsilon \times 10^{-2}$	$\Delta R/R_0$	$\varepsilon \times 10^{-2}$	$\Delta R/R_0$
2	Xm 1, 2, 3	0,48	-0,10	0,48	-0,11	0,48	-0,12	0,48	-0,12
	12, 13, 14	0,47	0,08	0,47	0,07	0,47	0,07	0,47	0,06
3	3, 7, 12	0,48	0,04	0,48	0,02	0,48	0,01	0,48	0,01
	6, 10, 13	0,48	-0,06	0,48	-0,07	0,48	-0,07	0,48	-0,07
4	5, 7, 12	0,55	0,02	0,55	0,01	0,55	0,01	0,55	0,01

Table 5. Gage factors (GF) and squared regression values (R^2) of established trend equations

Rod Type	**GF**		**Xm**	**R^2**		**Response**
	1*	**2***		**1**	**2**	
2	0,58100	0,46312	0,52206	0,98965	0,99602	Positive
	0,09554	0,12951	0,11253	0,95651	0,97279	
	-0,28869	-0,34058	-0,31464	0,99218	0,97288	Negative
3	0,16095	0,15806	0,15951	0,98918	0,99668	Positive
	0,03847	0,07018	0,05433	0,98339	0,98172	
	-0,14084	-0,16380	-0,15232	0,93925	0,96892	Negative
4	0,03581	0,03702	0,03415	0,97118	0,99305	Positive
	-0,07451	-0,07684	-0,07568	0,99448	0,98699	Negative

2.4 Braided Composite Rod and Steel Rebar Tensile Performance Comparison

Considering the different composite rods tensile strength, extension at failure, tensile strength at 0,2% strain and modulus of elasticity, some conclusions can be withdrawl.

BCR 4 presents the most interesting tensile performance while BCR 1 presents the less interesting one, although rod 1 presents the highest reinforcement fibre mass fraction.

BCR 2 and 7, presenting the same amount of E-glass and carbon fibres, presents significantly different tensile behaviour, mainly due to the reinforcement fibre mass fraction. Although rod 7 presents also HT polyethylene fibres, its fibre mass fraction is lower than in rod 2.

For composite rods 3, 6 and 5, with the same amount of E-glass and carbon fibres, the presence and increasing of HT polyethylene fibre, as well as the increase of the fibre mass fraction, promotes an increasing of the rod tensile performance.

Although the tensile performance of the braided composite rods is influenced by the reinforcement fibre mass fraction, one can conclude that the type of reinforcement fibre has a significantly higher influence.

When compared to the steel rebars currently used in the construction industry, composite rods reinforced by carbon, glass and polyethylene fibres present higher tensile strength. Current Portuguese steel rebars, A235NL, A400NR/ER and A500NR/ER have values of tensile strength of 360 MPa, 460 MPa, and 550 MPa, respectively. BCR 1 is the only composite rod that presents tensile strength lower than 550MPa. Even though the tensile strength of E-glass, carbon and HT polyethylene braided composite rods is higher than that of steel rebars.

However, composite rods have a lower modulus of elasticity when compared to that of steel rebars, 210 GPa.

3 Conclusions

The braided composite rods diameter varies due to the core reinforcement fibres used and to the resin mass fraction. There is no relationship between the rod diameter and the mass fraction of the reinforcement fibres. Braided composite rod, reinforced by 77% E-glass and 23% carbon fibres, presents the highest tensile strength. The lowest tensile strength is presented by the composite rod reinforced by 100% E-glass fibre. Analysing the extension at failure parameter, composite rod reinforced by 52% E-glass, 46% carbon and 3% HT polyethylene presents the lowest extension at failure. Once again, the composite rod reinforced by 100% E-glass fibre presents the highest value.

Braided composite rod reinforced by 100% carbon fibre presents the highest yield stress and, therefore, the highest modulus of elasticity. Composite rod reinforced by 100% E-glass fibre presents the lowest values in both parameters.

The BCRs that present the best tensile performance are those who present the lowest amount of E-glass fibre. Among the rods with the same amount of E-glass and carbon fibres, the composite rod with highest percentage of HT polyethylene presents highest tensile performance. The type of reinforcement fibre used has higher influence than the fibre mass fraction in the tensile performance of the FRP rods.

In what concerns the sensing performance of braided reinforced rods using glass and carbon, it can be concluded that all three types of BCR used can stand as a self-sensing material. The electrical contact set-up was effective in the purpose of resistance stabilization and measurent.Two types of responces were obtained by the BCR. Positive GF, in the case of the carbon fibreplaced in the area subjected to tensile and negative GF, in the case of the carbon fibre placed in the compressive side of the rod. Furthermore, the GF increased with decreasing carbon fiber content. The most reliable monitoring behavior was given by BCR 2 (77% glass, 23% carbon) with the smallest carbon fiber content.

References

[1] Muto, N., Arai, Y., Shin, S.G., Matsubara, H., Yanagida, H., Sugita, M., Naka-tsuji, T.: Hybrid composites with self-diagnosing function for preventing fatal fracture. Comp. Sci. and Tech. 61, 875–883 (2001)

[2] Chung, D.D.L.: Self-monitoring structural materials. Composite Materials Re-search Laboratory, State University of New York, Buffalo, NY 14260-4400, USA (1997)

[3] Bakis, C.E., Nanni, A., Terosky, J.A., Koehler, S.W.: Self –monitoring, pseudo – ductile, hybrid FRP reinforcement rods for concrete applications. Comp. Sci. and Tech. 61, 815–823 (2001)

[4] Lees, J.M.: Fibre.reinforced polymers in reinforced and prestressed concrete applications: moving forward. Prog. Struct. Eng. Mater. 3, 122–131 (2001)

[5] Soebroto, H.B., Pastore, C.M., Ko, F.K.: Engineering design of braided structural fibre-glass composite. In: 6th Annual Conference, Advanced Composites, Structural Composites: Design and Processing Technology, Detroit (1990)

[6] Fangueiro, R., Sousa, G., Araújo, M., Gonilho Pereira, C., Jalali, S.: Core reinforced composite armour as a substitute to steel in concrete reinforcement. In: International Symposium Polymers in Concrete – ISPIC 2006, Universidade do Minho, Guimarães, Portugal, April 2-4 (2006)

[7] Kang, I., et al.: Introduction to carbon nanotube and nanofiber smart materials. Composites: part B 37, 382–394 (2006)

[8] Li, X., Levy, C., Elaadil, L.: Multiwalled carbon nanotube film for strain sensing. Nanotechnology 19, 045501 (7pp) (2008)

Systems Assurance, Complexity and Emergence: The Need for a Systems Based Approach

Ali Hessami[2] and Nicos Karcanias[1]

[1] Systems & Control Centre, City University
[2] IEEE SMC Systems Safety & Security Technical Committee

Abstract. The complexity of modern products, systems and processes makes the task to identify, characterise and provide sufficient assurance about the desirable properties a major challenge. Stakeholders also, demand a degree of enhanced confidence about the absence of undesirable properties with a potential to cause harm or loss. The paper develops a framework of seven fundamental facets of performance as an ontology for emergent behavioural properties and a separate framework for the emergent structural properties of complex systems. The emergent behavioural aspects are explored and we develop a systems framework for assurance based on an Assessment and Management paradigm each comprising a number of principles and processes. The key argument advanced is that in the face of complexity and incessant change, enhanced confidence in the achievement of desirable and avoidance of undesirable properties requires a systems approach empowered by suitable modelling and relevant diagnostic tools explaining the nature of emergent properties. The principal focus of this paper is on safety, security and sustainability emergent behavioural (performance) aspects of complex products, systems and processes.

Keywords: Safety, Security, Complexity Sustainability, Assurance, Systems Approach.

1 Introduction

Amongst many challenges arising from the pervasive complexity in most modern products, systems and processes is the necessity to identify, characterise and provide sufficient assurance about the desirable properties. Alongside this, most key stakeholders, specifically the regulators and end users, demand a similar degree of enhanced confidence about the absence of undesirable properties often with a potential to cause harm or loss, for such products, systems or processes. We develop and propose a framework of seven fundamental facets of performance as an ontology for emergent behavioural properties and a separate framework for the emergent structural properties in complex and/or large scale system of systems. Understanding and managing complexity, as well as characterising structure are central to this work. The need for conceptualisation, analysis, assessment and enhanced confidence in the properties of complex systems, specifically the emergent behavioural aspects is subsequently explored where we develop and propose a systems framework for

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 202–215, 2010.

assurance based on an Assessment and Management paradigm each comprising a number of principles and processes. The key argument advanced is that in the face of complexity and incessant change, enhanced confidence in the achievement of desirable and avoidance of undesirable properties itself requires a systems approach, supported by appropriate modelling tools and diagnostics. These are needed to understand the nature of emergent properties as features of aggregation in complex processes and thus help us to avoid making erroneous decisions with costly and sometimes irreversible consequences. The principal focus of this paper is on safety, security and sustainability emergent behavioural (performance) aspects of complex products, systems and processes, but the framework has more general validity.

2 Complexity and Emergent Properties

Complex Systems is the term that emerges in many disciplines and domains and has many interpretations, implications and associated problems. The features of a specific domain characterise the dominant forms of complexity associated with the problem. A very significant class of complexity issues is linked to the multidimensionality of views of a system and in particular the design and operation in the case of industrial systems. Figure (1) describes the basic ***system shell*** and this indicates the multi-view of the system that is linked to:

(i) Physical Process Dimension
(ii) Signals, Operations Dimension
(iii) Data, IT, Software Dimension
(iv) Embedding in the Environment

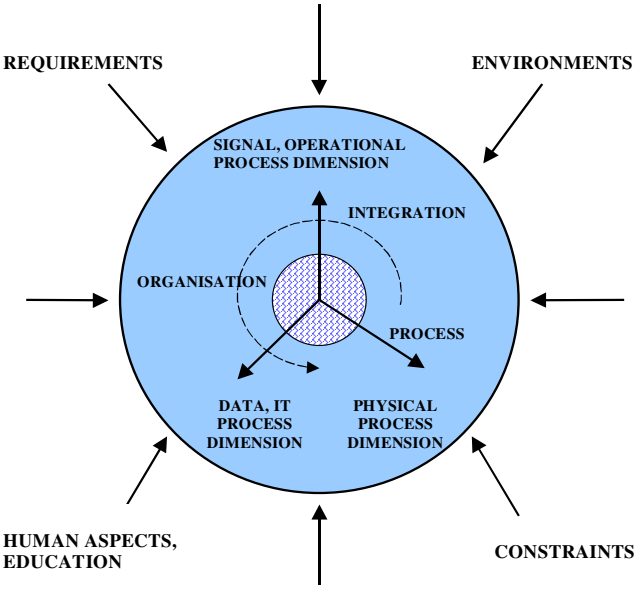


Fig. 1. Basic System Shell of Manufacturing Integration

The distinguishing features of this viewing of the system are the close links between modelling, system structure, system organisation, measurement, structures information, control, decision, management and resulting system properties and such a study requires a systems framework.

In this paper our interest is focused on aspects of systems performance. The performance of complex systems is a measure of their utility, output and perceived emergent properties and central issues to this study are: **(i)** Characterisation and Management of System Complexity; **(ii)** Emergent structural and non-structural properties; **(iii)** Emergent behavioural properties. Problem complexity is manifested in many different ways which include:

- **(a)** Lack of knowledge, or difficulties in characterising the behaviour of the basic process (*Unit Behavioural Complexity*).
- **(b)** Complexity of computational engines (*Computational Complexity*).
- **(c)** Difficulties in characterising the interconnection topology of sub-processes and/or variability, uncertainty of this topology during the system lifecycle (*Interconnection Topology Complexity*).
- **(d)** Large scale dimensionality (*Large Scale Complexity*)
- **(e)** Heterogeneous nature of sub-processes, resulting in hybrid forms of behaviour (*Hybrid Behavioural Complexity*).
- **(f)** Organisational alternatives for the functioning, information and decision making (control) structures in respond to goals and operational requirements (*Organisational Complexity*).
- **(g)** Variability and/or uncertainty on the system's environment during the lifecycle requiring flexibility in organisation (*Lifecycle Complexity*).
- **(h)** Uncertainty in describing the embedding of the system in its environment (*Environment Embedding Complexity*)

Emergent properties refer to aggregate aspects of behaviour of the system properties which are frequently linked to specific metrics defined by the system variables. The emergent behavioural properties of complex systems comprise an ontology of seven often context sensitive facets namely: **(1)** Technical functionality; **(2)** Cost; **(3)** Environmental behaviours & Sustainability; **(4)** Reliability, Availability, Maintainability; **(5)** Safety & Security; **(6)** Quality; **(7)** Perceived Value. Such properties are reasonably distinct and often inter-related, thus posing a major challenge to designers, to arrive at optimum solutions which satisfy stakeholders' expectations on each dimension. The evaluation of their degree of presence, or absence and the nature of interrelationships between them is a challenge that frequently depends on the nature of the specific system. A key distinction between these emergent properties is the fact that apart from safety, security and environmental performance, which are subject to a regulatory framework, the desirable level for the rest of these properties e.g. cost, reliability, quality etc. are left to the discretion of the duty holders and market forces. This therefore creates a legal compliance issue for attaining and assuring certain characteristics as well as deliver the corporate social responsibility.

The key differentiation between *safety* and *security* performance is: safety is freedom from harm to people caused by unintentional or random/systematic errors and failures of a product, process, system or mission whilst security is freedom from

loss caused by deliberate acts perpetrated by people. Therefore security is principally characterised by intent and nature of causation as opposed to strictly being an output performance indicator reflecting degrees of loss or gain. Like safety performance, security of a system is mainly measured probabilistically in terms of risk due to inherent uncertainties.

The security of systems is often forecast and measured in terms of perceived or real threats and vulnerabilities and not in terms of consequential risk of harm and loss. The threat is often an external source of malicious intent whereas vulnerability is an inherent flaw/dysfunction in a system making it prone to external and sometimes internal threats. There's a lack of systemic approach in identification, assessment and management of such risks in most enterprises and endeavors. This paper develops a systemic framework for assurance of safety and security in complex systems whilst proposing an innovative set of performance criteria for these critical facets of emergence/performance. We further endeavour to develop the case for a unified approach to emergence, assessment and management of emergent properties in complex products, processes, systems and undertakings. To this end, we propose sustainability provides a candidate unifying framework in the sense that, from a holistic perspective, any product, process, system or undertaking which lacks the right blend of desirable emergent properties such as safety, quality, reliability, affordability, environmental friendliness, social acceptance etc. can be viewed as unsustainable.

3 Systems Safety and Security, the Fundamentals

A. System Safety Concepts

The classical view of safety performance in hard and soft systems [5] is often biased towards historical accidents and often feeble post mortem attempts at understanding the causation and prevention . This deficient and primordial paradigm is challenged on the grounds that:

- Same accident may arise from a multiplicity of different causative factors;
- Accident investigations are predominately driven by legal imperatives and the need for finding a responsible person/body as opposed to the systemic understanding of the underlying root causes;
- Increasing pace of change, innovation and complexity in modern systems creates opportunities for new forms of accidents as yet un-encountered;
- The social, legal and organizational costs linked with accidents are increasing due to public awareness, regulation and the litigation process.

It is argued therefore that allowing accidents to happen and the subsequent often inconclusive and feeble attempts at investigation and learning is tantamount to negligence and admission of failure in the face of challenges and risks faced. A new advanced paradigm based on credible and objective scientific principles is needed to counter the formidable risks posed by modern complex undertakings.

(1) The Systems Approach to Safety

In view of the major shortcomings of the classical accident focused approach cited above, the systems approach to specification, realisation and management of safe and secure systems is founded on the identification of hazardous states, generally precursors to accidents. This generates a deeper insight in complex behaviours and can expose a vast array of faults, errors, failures and vulnerabilities which may lead to the realisation of hazardous states. Likewise, a hazard focused approach provides the opportunity to objectively scrutinise the potential escalation scenarios associated with a hazard and devise solutions to detect, contain, control or mitigate the broad range of accidents which may arise from such states in a system.

In sharp contrast to the reactive learning from accidents, the systems approach to safety assurance focuses on empirical as well as creative identification of hazards. Once a suite of key hazardous states are identified and ranked, it explores their causes, random or systematic [7], scrutinises their escalation scenarios and devises risk control and mitigation strategies [1]. Crucial for this is the need for a general systems framework that defines the relevant states.

(2) The Need for System Safety Metrics

Safety is a human focused concept reflecting the degree of freedom from unacceptable harm to people. Paradoxically, it is often measured by its absence for example, the safety of products, processes, and systems is regularly quoted in terms of risk of harm they may cause to specific groups as opposed to the expected duration of harm free operation akin to reliability! The other fallacy is to forecast the safety of a complex system principally based on the empirical or past performance of similar systems, a notion which relates to random rather than systematic causes of hazards naively assuming that the future is a simple (linear) evolution of the past. The irony being that most modern and complex systems principally suffer from the systematic errors due to pervasive incorporation of embedded intelligence.

Safety is predominately measured in terms of risk which is a forecast comprising the likelihood/frequency of an accident and the degree of loss that it may entail. This poses a challenge to duty holders or system designers who find it difficult to relate the faults and failures of their products or systems to likely injuries and fatalities to the end users. To this end, some system standards [7] have advocated hazard rates as a direct measure of system safety, leading to the classification of system's safety properties in terms of Safety Integrity Level (SIL). The SIL concept which has a widespread following in industry is more akin to a reliability perspective and is a non-systemic convention without much regard to the consequences of the dangerous failures [9]. It simply considers a range of potential functional failures in probability or frequency that are undesirable and considered to be dangerous, lacking a systemic appreciation of the real world implications of such failures. They are just called dangerous without a unit declared for danger! This is a far cry from science in safety even though exceedingly small numbers such as 1E-09 are employed as acceptable dangerous hourly functional failure rates for high dependability systems.

Some sector standards, strangely derivatives of the IEC system standard [7], such as those for safety critical transport [8] advocate Tolerable Hazard Rates (THR), taking into account a total systemic perspective and the notion of tolerability of risk. We need systemic metrics which go beyond failure and take into account exposure of

various groups at risk and the potential escalation scenarios and tolerability criteria [17]. The THR concept, principally reliant on historical performance of systems, goes a fair way towards this ideal but fails to explicitly address all requisite factors in a single metric. There's a need for a portfolio of systemic lead as well as lag indicators for safety, security and sustainability of complex systems.

B. System Security Concepts

Unlike safety, security has many different interpretations and implications for its stakeholders. From a systems perspective, security is lack of susceptibility to malicious intent which may comprise; **(i)** Vandalism; **(ii)** Sabotage; **(iii)** Theft and fraudulent gain; **(iv)** Terrorism; or a combination thereof. Security or lack of it is principally characterized by the intent on causing harm and therefore, it is a mostly human focused issue. However, in the cybernetics domain, this may become a concern between autonomous intelligent systems without direct human intervention [6].

(1) The Systems Approach to Security

There are two fundamental facets to security of a general system. The extrinsic dimension or driver is *threat*, characterized by the real or perceived existence of people or systems with intention to cause harm and loss. The intrinsic dimension or counterpart is *vulnerability*. Whilst threats are diverse and unlikely to be fully forecast, anticipated or controlled, vulnerabilities are characteristics of a general system, which arise from lack of awareness to potential for harm from threats in the larger environment of operation. Frequently, vulnerability may be characterized as a structural system property linked to interconnection topology, or some system functionality with a critical role, or linked to external to the system factors (external influences). Defining system vulnerability in concrete terms requires diagnostics and an appropriate methodology.

The main thrust of systems security assurance rests upon systematic identification of key vulnerabilities, analysis of the causations and potential escalation scenarios and evaluation of pertinent risks. This is followed by proactive development of elimination or control strategies for major vulnerabilities and identification of detection, containment or mitigation solutions in the event of realisation of threats. However, similarly to the systems safety related precursors (hazards), vulnerabilities, seen as aspects of a system's architecture or operation are mostly a concern at the system boundary. An elaboration of this may lead to the consideration of internal and external threats and vulnerabilities with major implications for systems security (beyond the scope of the current debate). In Systems of systems (SoS), or large open systems with significant vulnerabilities, security is often assured through focus on threats rather than vulnerabilities.

(2)The Need for System Security Metrics

Bearing in mind the extrinsic and intrinsic facets, it is instructive to identify, quantify and treat threats and vulnerabilities collectively to ensure completeness and coverage of key concerns. Threat as an extrinsic measure for a system's security is generally classed into a number of distinct levels. The US Department of Homeland Security defines five Threat Conditions, each identified by a description and corresponding colour. From lowest to highest, the levels and colours are: **(a)** Low = Green;

(b) Guarded = Blue; **(c)** Elevated = Yellow; **(d)** High = Orange; **(e)** Severe = Red. However, these are principally threat criteria relating to terrorism, whereas risk includes both the probability of an attack occurring and its potential losses.

In a similar manner to the threats, metrics are called for systems vulnerabilities since these render a system susceptible to damage and harm, even in the absence of malicious intent at the outset. Even though the safety concept of SIL is not truly indicative of safety properties of a complex system [9], it is more appropriate for measurement of vulnerability since this is an intrinsic (architectural, compositional and operational) system property. A credible metric for system's vulnerability would provide an objective measure of its resilience against potential threats. This could be a *System Resilience Index* which needs to be elaborated and quantified for various classes of vulnerability.

C. System Sustainability Concepts

(1) The Systems Approach to Sustainability

Sustainability is a high level emergent system property that expresses the ability of the system to survive and continue to function according to the original goals set for its operation. It is thus related to :

(i) Robustness of the system behaviour to external disturbances ;
(ii) Ability to overcome threats that may have catastrophic consequences by demonstrating capabilities to survive and achieve the central goal ;
(iii) Adaptability by demonstrating capability to reorgonise its control and information structures after some catastrophic events, or changes in the operational goals of the system due to changes in the market ;
(iv) Potential for the system to evolve in a continuously changing environment of goals, specifications and constraints.

In principle, apart from survivability and résilience attributes, sustainability possesses social, economic and enviornmental dimensions as well, making it a complex composite property in its own right. It is clear therefore that the basic concepts required to define sustainability are themselves emergent system properties and it is this that makes sustainability a higher level emergent property.

(2) The Need for System Sustainability Metrics

Defining *sustainability* as an emergent higher level, composite property implies the need to:

(i) Identify the constituent primitive emergent properties. **(ii)** Develop diagnostics for characterising and evaluating these properties. **(iii)** Develop a coneptual system framework expressing sustainability as composition, aggregation of the constituent emergent properties. **(iv)** Develop a meta-model expressing this aggregation and enabling the evaluation-estimation of sustainability.

Developing sustainability metrics is very challenging and requires addressing all previous issues .The difficulties are due to the characterisation of primitive emergent properties in a quantifiable way, and expressing their composition in a form that supports development of composite metrics.

4 Systems Safety, Security and Sustainability Assurance : The Framework

We propose two complementary and advanced sets of systems principles and processes as the underpinning backbone to tackling the challenges of safety, security and potentially sustainability. Taking a life-cycle perspective [12] these comprise I & III below;

■ **Assessment:** This comprises recognising the need, defining the system, specifying and identifying/understanding of key properties, behaviours, hazards and vulnerabilities, evaluating and assessing expected impact;
■ **Realisation:** This is ultimately aimed realising the desirable properties and achieving the desired performance in the form of product, process, system, mission or undertaking;
■ **Management:** this comprises taking the outcome of assessment and realisation into consideration and ensuring deployment, delivery of requisite performance, continued monitoring and control through a responsive and holistic suite of strategies and actions.

Whilst Realisation is specific to a given domain and context, the Assessment and Management aspects as a suite of principles constitute a meta-knowledge framework which can be abstracted and developed for almost universal application across many domains and disciplines. The systemic framework of assessment and management is equally applicable and effective within the context of desirable as well as undesirable properties of systems. This is contrary to the current conventional wisdom where specification, delivery and continual monitoring of desirable aspects of performance is regarded as an essentially domain expertise where as the undesirable and unintended emergent properties (hazards and vulnerabilities) are the forte of so called risk management. The +Safe3 extension [11] to the renowned CMMi model [14] also distinguishes between Safety Engineering & Safety Management, which are mainly synonymous with Risk Assessment and Risk Management advocated here.

Whilst presented as a dual and complementary suite of principles and processes, assessment and management are iterative and systemic in the sense that processes inherent in the management framework employ assessment activities at requisite points to support judicious decision making and ensuring optimal performance. These are collectively referred to as Systems Assurance and labeled as Surety Framework in this paper.

A. Risk Assessment

This key facet of Surety framework depicted in Fig. 2 is proposed as a backbone to the identification, specification, evaluation and assessment of the undesirable events or properties adversely affecting technical functionality, cost, reliability, safety, quality etc. The risk assessment process [13] comprises seven systemic aspects such as: **(a)** Hazard Identification; **(b)** Causal Analysis; **(c)** Consequence Analysis; **(d)** Loss Analysis; **(e)** Options Analysis; **(f)** Impact Analysis; **(g)** Demonstration of Compliance.

The risk assessment process, aims to enhance the systemic understanding of the key issues and it is not an end in itself. Assessment process generates transparency and awareness of real and potential issues thus empowering the duty holders to take appropriate actions and make the transition from fire fighting and reactivity to anticipation and proactivity.

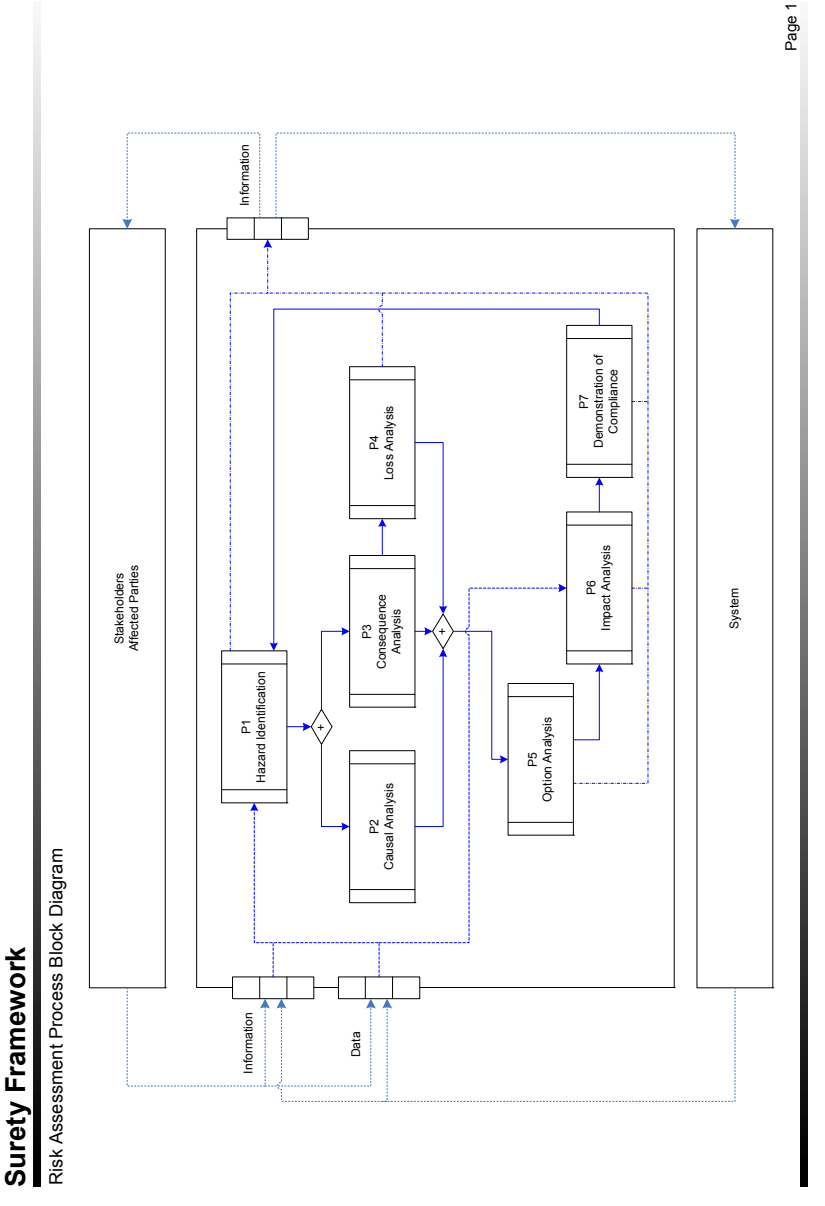


Fig. 2. Risk Assessment approach in the Surety Framework

B. Risk Management

A holistic and systemic approach to assurance of safety and security is developed and proposed in a major paper [4]. The paper elaborates seven principles which have to be collectively fulfilled before sufficient assurance is gained and maintained in the desirable safety and security properties of a general system. This complementary aspect of assurance within the Surety Framework comprises an advanced and systematic approach to developing, sustaining, enhancing and managing the so called downside events and properties associated with any system. Risk management builds upon the outcome of systematic assessment and ensures the identified and prioritized risks are eliminated, mitigated or continually controlled in a comprehensive and responsive manner. The risk management process is depicted in Fig. 3.

The proposed systems set of principles demands a detailed scrutiny of the problem domain, as the key stage in safety/security assurance followed by a number of complementary and value added activities. The principles underpinning the systemic management of safety and security are;

(1) Proactivity; **(2)** Prevention; **(3)** Protection & Containment; **(4)** Preparedness & Response; **(5)** Recovery & Restoration; **(6)** Organization & Learning; **(7)** Continual Enhancement.

These principles are detailed in [4]. However, the suite of seven principles is equally applicable to systems in which, in view of the complexity or novelty, assurance is mainly derived from the quality of the process and competencies of those involved.

C. Application of the Framework

The systemic framework of assessment and management proposed here is applicable to the attainment, maintenance and enhancement of three key and increasingly regulated aspects of safety, security and the environmental performance/sustainability of systems.

Nano-technology poses a modern and innovative domain where the safety and indeed security and the environmental implications of its products and offerings are largely unknown. An illustrative case involves the marketing of cosmetics containing nano-particles [10]. Because of their far smaller size, these particles are absorbed deeper into epidermis, dermis, cells and eventually into the blood stream of the users. The significant uncertainty on the risks has led to calls from the UK Royal Society and the US Federal Drug Administration (FDA) for research into the likely effects. The cosmetics industry considers nano-particles a “hot technology” with lots of intriguing applications, allocating vast sums to research into nano-technology. The FDA maintains that urgent research is called for due to the paucity of the knowledge on the effects of the nano-particles when they enter cells in the human body. A systemic framework constitutes a potent weapon in the face of such huge uncertainties with major implications for the human society at large.

The seven underpinning principles for risk management can be mapped to the requirements of any domain at any level of abstraction or details namely: **(i)** Industry / Sector; **(ii)** Corporate / Organization; **(iii)** Division / Team; **(iv)** Project / Product; **(v)** Mission. The scalable architecture for application of the proposed surety framework at society/corporate) and system/product levels would entail:

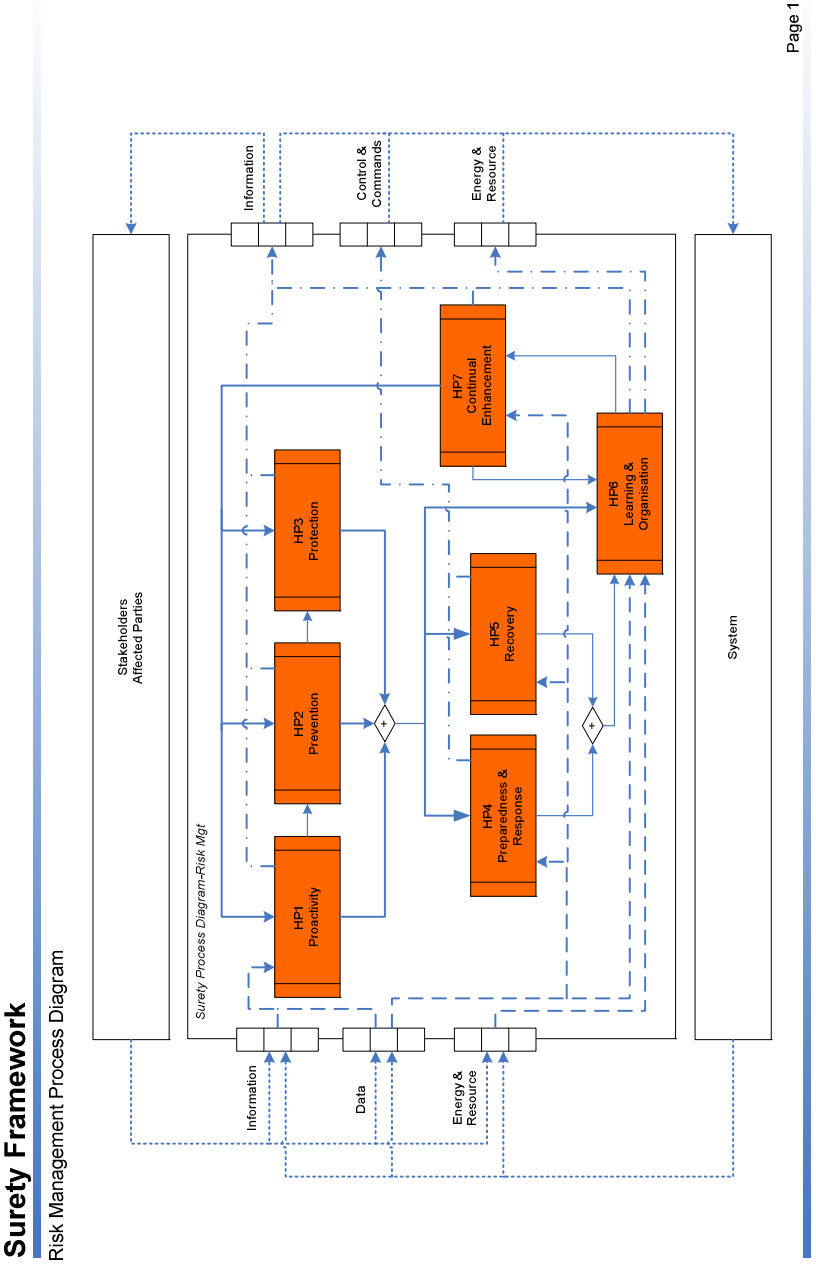


Fig. 3. Systems Risk Management approach in the Surety framework

(a) Identification of key influencing factors for each of the seven principles and generation of a hierarchical model for such factors depicting their roles and relationships [2];
(b) Assessment and quantification of these models and generation of an overall numerical index for each principle in the framework [3];
(c) Generation of a combined figure of merit (System Integrity and Resilience Index-SIRI) for the system, based on the seven indices derived for each principle.

Such indices can be benchmarked against desirable or tolerable levels of safety, security and environmental performance thus providing a reference for the optimal assurance under each principle as well as the whole framework applied to a system. This generates a focused and responsive system for attainment, management and continual enhancement of safety and security properties at the pertinent application level.

5 Conclusions

Amongst the seven key facets of a system's performance cited earlier, the safety, security and the environmental/global aspects are increasingly regulated by governments [17, 19]. This is partly driven by the gradual enhancement in the quality of life and public's awareness and demand for a more socially responsible stance by duty holders; private and public corporations, service providers and the suppliers. One of the striking observations in the fields of safety, security and environmental assurance is the overt reliance on often parochial technical solutions at the expense of a systemic and holistic understanding of the key issues and domain requirements.

Cybernetic systems driven by complexity, novelty and increasing pace of change and progression pose a challenge in safety and security if not environmental assurance due to inherent uncertainties. In such settings, the adoption of a systemic framework of universal principles assists with enhanced confidence in emergent properties where otherwise significant uncertainties prevail. The proper development of the field requires a suitable abstract systems framework that can explain and provide model based tools and diagnostics for emergent system properties. This is crucial for the development of metrics that can characterize primitive and composite emergent properties. Metrics may provide characterization of such properties. Linking emergent properties to system structure is critical, if we are to address issues of re-engineering of systems and processes aiming for development of systems with improved desirable properties, or reduced risks. Engineering/reengineering for improved systems assurance is an area where future research has to develop. Such efforts, however, require an appropriate systems framework [15], [16] that can support analysis and design by following paths similar to those deployed for hard systems.

We have developed and proposed an integrated framework comprising assessment and management paradigms collectively labeled as Surety. However, whilst the current focus has been the avoidance or minimization of risks, Surety framework additionally encompasses performance enhancement and optimization not addressed here. Such systemic assurance frameworks are instrumental in holistic identification, classification and treatment of critical issues (hazards and vulnerabilities) and the

specification/adoption of pertinent solutions. Founded in systems theory and embodying a significant structural, empirical and scientific knowledge, they also assist with the evaluation of the effectiveness of the risk control options whilst exposing gaps in the overall landscape and strategy. In view of the synergies between safety and security facets of performance, adoption of one integrated framework would result in savings on time and effort whilst optimising investment in equipment and systems. They are the most potent weapon in the face of epistemic uncertainty.

Beyond this intermediate development, we find the landscape of systems covered by many disparate specialisations, parochial expertise and lack of holism in approaching emergence, risk and opportunity in any complex setting. To this end, we will explore the concept of sustainability as a potent unifying umbrella for all emergent properties in a product, process, system or undertaking. Failing to find such systems based unification will result in more chaos and confusion about emergence and coherent engineering of emergent properties since each facet requires extensive expertise and competence to assure. Further research will be needed to develop a rational and holistic case and framework for this unification. Systems paradigm will provide the engine for this profound understanding.

6 Nomenclature

Assurance: Increasing confidence and certainty
Gain: Lives saved, improvements made, damages prevented or avoided in the natural habitat or benefits accrued to a business /society or a combination thereof. The expected value of a future benefit.
Hazard: Object, state or condition which in the absence of adequate detection or containment could lead to an accident.
Health: Soundness of body and mind, freedom from illness
Loss: Physical harm to people, detriment to a business/society or damage/destruction of the natural habitat or a combination thereof.
Reward: A forecast for a desirable event entailing a gain.
Risk: A forecast for an accident or loss. The expected value of a future loss.
System: A (purposeful) composite of inter-related parts / constituents with discernible collective output(s) or emergent property(ies) not manifested by any of the elements.
Safety: Freedom of people from (physical) harm.
Security: Freedom from vulnerability or loss caused by deliberate and malicious acts.
Sustainability: A blend of social, economic and environmental considerations which render a product, system or undertaking viable and continually optimal.
Systems Assurance: The art, science and technology of ensuring and demonstrating that a system is likely to achieve its objectives without engendering unacceptable levels of loss.
Systems Safety: The art, science and technology of ensuring and demonstrating that a system is not likely to lead to unacceptable levels of (physical) harm to people.
Systems Security: The art, science and technology of ensuring and demonstrating that a system is not likely to be vulnerable to malicious deliberate acts aimed at engendering unacceptable levels of loss.
Vulnerability: Susceptibility to injury, fatality or loss.
Welfare: Well being and quality of life for individuals and the society.

References

[1] Hessami, A.G.: Safety Assurance, A Systems Paradigm. Hazard Prevention Journal of System Safety Society 35(3), 8–13 (1999) (third quarter)

[2] Hessami, A.G.: Risk, A Missed Opportunity. Risk and Continuity Journal 2, 17–26 (1999)

[3] Hunter, A., Hessami, A.G.: Formalization of Weighted Factors Analysis. Knowledge-Based Systems (2002)

[4] Hessami, A.G.: A Systems Framework for Safety & Security, The Holistic Paradigm. Systems Engineering-The Journal of the International Council on Systems Engineering 7(2), 99–112 (2004)

[5] Waring, A.: Practical Systems Thinking. International Thomson Business Press (1996) ISBN 0-412-71750-6

[6] Clarke, R.: Asimov's Laws of Robotics, Implications for Information Technology. IEEE Computer, 53–61, 27, 57–66 (December 26, 1993–January 1, 1994)

[7] IEC 61508: Functional safety of electrical/electronic/programmable electronic safety-related systems - International Electrotechnical Commission (January 20, 2005)

[8] CENELEC: European Standard EN50129 Railway Applications – Communications, Signalling and Processing Systems – Safety Related Electronic Systems for Signalling (February 2003)

[9] Hessami, A.G.: Risk Management a Systems Paradigm. Systems Engineering-The Journal of the International Council on Systems Engineering 2(3), 156–167 (1999)

[10] Safety Fears over 'nano' anti-ageing Cosmetics. The Sunday Times (July 17, 2005)

[11] +Safe Version 1.2, A Safety Extension to CMMi-DEV Version 1.2, Defence Materials Organisation, Australian Department of Defence (March 2007)

[12] ISO/IEC15288, System Life Cycle Processes - ISO/IEC (October 2002)

[13] Engineering Safety Management Issue 3 (Yellow Book III), vol. 1&2, Fundamentals and Guidance, Railtrack PLC UK (January 2000) ISBN 0 9537595 0 4

[14] Chrissis, M.B., Konrad, M., Shrun, S.: Guideline for Process Integration and Product Improvement, CMMI Second Edition (January 2007) ISBN 0321279670

[15] Karcanias, N.: System concepts for General Processes: Specification of a new Framework. Systems Research Centre Report 2003, SRCRep-06-03/1, City University (2003)

[16] Karcanias, N.: Structure evolving systems and control in integrated design. IFAC Annual Reviews in Control 32(2), 161–182 (2008), doi:10.1016/j.arcontrol.2008.07.004

A Review on Sustainability Models

Amin Hosseinian Far, Elias Pimenidis, Hamid Jahankhani,
and D.C. Wijeyesekera

School of Computing, IT and Engineering
University of East London, UK
{Hamid.Jahankhani,E.Pimenidis,amin,chitral}@uel.ac.uk

Abstract. The apprehensions around the climate change and energy crisis have led to the emergence of the need for sustainability analysis. The subject of sustainability is now of wider inclusivity as it can be applied to almost any field of study. There are different approaches to model sustainability; however the systems approach is the focus of consideration in this paper. The proposed model for sustainability incorporates the use of neural networks to develop a complex adaptive system. The complexity of the system is simplified using influence diagrams or knowledge management techniques.

Keywords: Knowledge Management System, Complex Adaptive System, Sustainability, Systems Theory.

1 Introduction

There are many different definitions for the term sustainability but what they all have in common is: it is referred to an ongoing process and can describe a systems' state. This means the sustainability of a system should be considered as a feature of the system. Here is a definition for the phrase: *"The ability to maintain a balance in a process or a state in a system, whether ecological, technological or social is currently known as sustainability."*[1]

It is feasible to have an illustrative model for sustainability. In addition of the concept description using illustration, it would be possible to apply the model into practical case studies.

2 Current Sustainability Models

There are some models for integrating the environmental criteria as well as the considering the business profitability towards implementation of a sustainable environmental friendly system. [2]. The ISO 14031 Standard considers the Environmental Performance Evaluation (EPE) as the implementation guideline and standard to model the new sustainable systems. The general model for implementation is the PLAN-DO-CHECK-ACT (PCDA) for business process representation as shown in figure 1. [3]

There are some models called corporate environment models that are illustrating the relationship between the company (including its management) and the environment and was firstly introduced in 1998.

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 216–222, 2010.

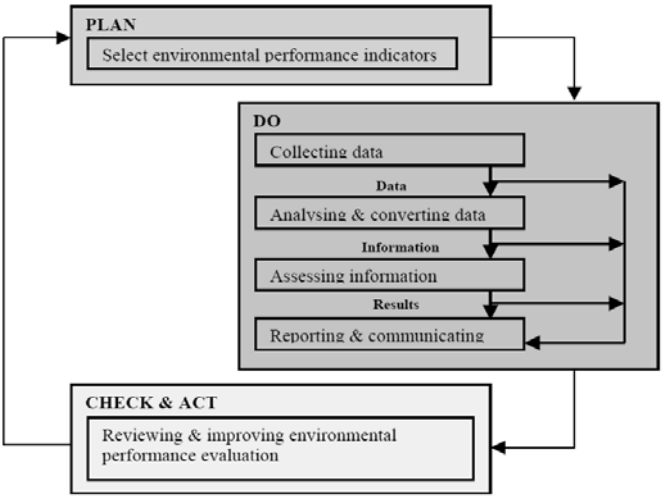


Fig. 1. Plan-Do-Check-Act model by Altech Environmental Consulting Ltd, ISO 14031

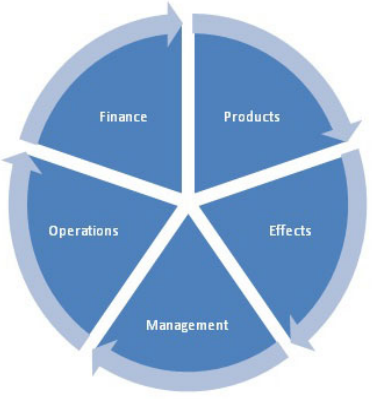


Fig. 2. Corporate environment modelling by Rikharsson normally figured in pentagon [4]

The product itself is set into two different categories: Product Balance [5] and Inventory stage [6]. The product balance would consider the environmental aspects of the system (or the company) itself, which might include the effects of the materials used and especially effects of the system on the environment. On the other hand the inventory stage considers the life cycle assessment of the system. The integration of each domain needs to have a calibration towards the sustainable system and the managers or stakeholders should decide upon the output of the overall system. Bennett *et al* 1999 suggests the following checklist for the final implemented information system that calibrates and integrates these different domains:

- Data support to environment managers
- Internal performance reporting
- External environmental reporting
- Support the implementation of ISO 14001 or other environmental management standards
- Other;

Although some of these items would be replaced based on the stakeholder's benefits.

The traditional and universal models for sustainability are the Russian doll model for sustainability by O' Riordon and also the John Elkington's Venn diagram:

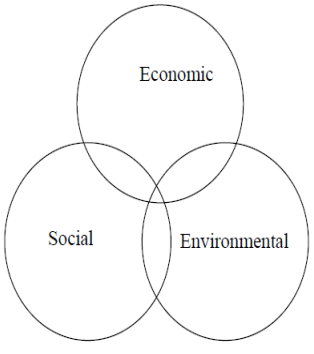


Fig. 3. O' Riordon Model [7]

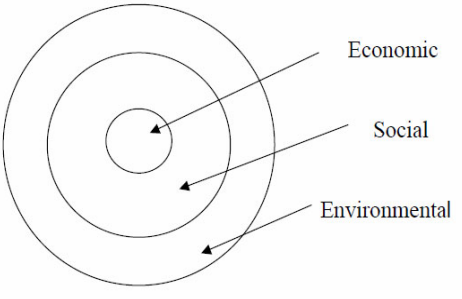


Fig. 4. Elkington Model [8]

What is noticeable in the newer model is that the economic domain has been placed as the core of the model. This is surrounded by social domain and these two are located within environmental outer concentric circle. There are many other different ways for illustrating sustainability in a model. It can be elucidated simply by qualitative analysis of all three domains. Some analysis has already been done for some of the domains using experiments and computer simulations. For instance, the economy and the environmental analysis can be done by using RetScreen tool which is an excel-based software using its macros and many imported libraries. This tool is implemented by Canadian Natural Resources and it seems to be one of the advanced tools in this field. But it does not consider the management domain and also the new policies and the new patents or technologies which are affecting the overall model. This and other practical tools are implemented in order to find some answers to questions in the real world. But still there is no extensive theoretical model that considers all the dynamicity and complexity of system sustainability.

One of the known approaches to model sustainability of systems in general is the system approach.

3 System Approach and Knowledge Management

The system approach would be one of the options for modeling sustainability where the relationships, axioms or rationale and domain can be easily defined in terms of systems or sub- systems [9]. The system approach looks at the project from the systems point of view. Principally a system needs to have the input, process and output domain. The inputs to a sustainable model or network are the knowledge that are affecting systems' sustainability. For instance economy factors, environment measure and also the social inputs are the major knowledge representing the sustainability sub systems. But each domain consists of many factors that each might have a different effect on the overall system. The knowledge management techniques would need to be implemented in order to identify the knowledge behind sustainability. After identification of the knowledge needed for the model, then the system techniques can be utilized to consider the calibration between the inputs or better to say the knowledge.

4 Boundary Dilemma

The system is sustainable while all the links and inter relations between different domains sustain for the defined period of time normally an infinite period [10]. If we consider a system such as a solar PV system as an example, the boundary could be very small or very large. The solar energy itself is sustainable, and a limited component has been considered in this case. All within the boundary we have is the energy. While we include more components within the boundary, the sustainability of the system is harder to predict and the current models cannot be mapped in more complex systems.

5 Knowledge Management- System's Theory, and Complex Adaptive System Modelling

Modeling a system affected by economic, social (policy making processes and standardization in this case), and environmental factors requires consideration of all components, parts and better to say domains interacting with the system. There is a proposal for policy translation and reconciliation process by Drew *et al* 2009 where the use of the neural networks is considered. [11]

The modeling may be considered from two points of view. One is System of Systems modeling approach and the other would be contemplation of complex systems.

As we look at our dilemma which is named as the boundary paradox, the only system tool that can easily model a system without considering the systems boundary (say an open system; although we do not want specifically consider an open system) would use the complex system theory. Furthermore Neural Networks and Artificial Intelligence can also be applied. The inputs of the system can be the nodes of the network. They should get leveled and more levels represent more complexity. More levels give more efficiency but it is more time consuming. [1] These networks have the ability to learn (which is needed in the case of new technologies and new policies), also the calibration between these different domains should be considered.

Moreover, there are other similar proposals which were forgotten at some point regarding computational sustainability modeling. For instance Clayton & Radcliffe in 1996 had the following idea:

"The question of sustainability affects most areas of human activity. It is intrinsically complex and multi-disciplinary. Sustainable policies have to adapt to new knowledge and changing circumstances. Understanding sustainability and ways of achieving it have to involve an understanding of complex adaptive systems and general systems theory - a rapidly developing new branch of social studies." [12]

Complex Systems can be divided into these types: Chaotic Systems which are affected by the butterfly effect and are not the focus of this work; Non-linear Systems which do not reflect the behavior and emergence of sub systems and are modeled non-linearly and the Complex Adaptive Systems that learn from the environment. The adaptive characteristics would be beneficial in reflection of policy making and previous market trends, [13].

There are some other modeling tools and methodologies for representation and analysis of complex systems. One of the famous analysis tools called Integrated Assessment (IA) is discussed below:

"Interdisciplinary process to integrate the knowledge from different disciplines and different stakeholder groups to evaluate a problem situation from synoptic and local perspective"

- IA should support decision processes
- IA should help to identify desirable and possible options "[14]

One of the widely used kits for the analysis of complex systems and using IA, is Influence Diagrams. This has been previously used in many areas and by well known organizations' projects such as NASA projects using the Analytica platform. The main reason for the use of Influence Diagrams is to simplify the complexities of the system. What are exactly modelled with Influence Diagrams are the concepts and domains and their influence on the other domains or even themselves.

6 Proposal

The proposed system approach to quantify sustainability can be considered as a network, receiving the knowledge representations as the input, and the quantitative sustainability factor as an output.

The above network should have the following characteristics:

- It has the adaptation capability; it learns from different inputs
- Adding the inputs will increase the complexity of the network; this might lead to an open system model
- Not so many feedbacks are considered for the network
- Leveling for a complex network would lead to delays on the output
- Many relationships within the network are non-linear

The relationships within the nodes should remain sustainable after the calibration is set, unless the system learning changes the path within the network

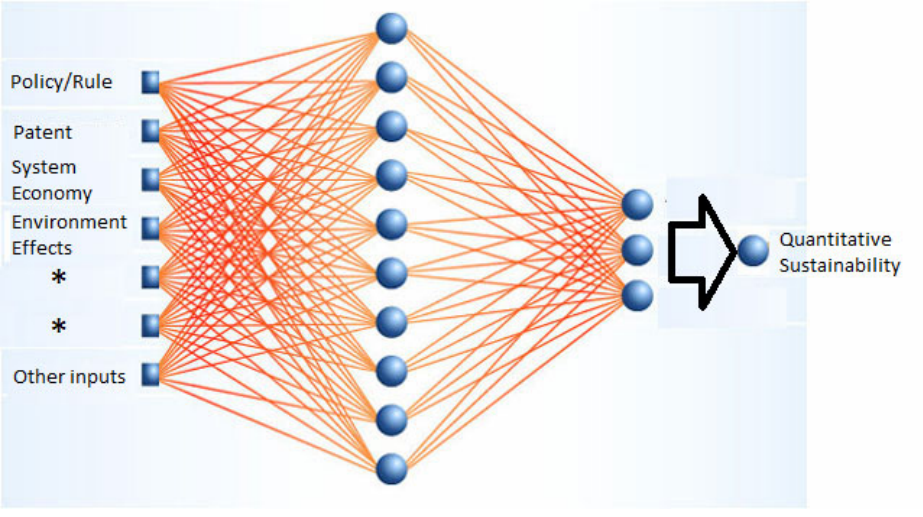


Fig. 5. Proposed model for sustainability using a neural network

7 Conclusion

Sustainability is considered as the ability to sustain, in addition it is the state or characteristic of a system, or process. There different ways for modelling the concept. Initially the model was introduced for specific knowledge domain; i.e. environment and banking system, but nowadays it can be applied to almost every domain.

System approach is a method for modelling sustainability. There many proposals considering different system techniques. Using knowledge management techniques would elaborate the knowledge behind sustainability. The knowledge then can be modelled in a neural network. The network would then demonstrate the relationships and axioms between each knowledge domain which are the inputs of the network.

References

1. Hessami, A.G., Jahankhani, H., Hsu, F.: A Systems Framework for Sustainability. In: Jahankhani, H., Hessami, A.G., Hsu, F. (eds.) ICGS3 2009. CCIS, vol. 45, pp. 76–94. Springer, Heidelberg (2009)
2. ISO: ISO 14031-Environmental Performance Evaluation – Guidelines. Canadian Standards Association (1999)
3. Putnam, D.: ISO 14031: Environmental Performance Evaluation. Confederation of Indian Industry Journal (2002)
4. Rikhardsson, P.M.: Information Systems for Corporate Environmental Management Accounting and Performance Mesurement. In: Bennett, M., James, P. (eds.) Sustainable Measures-Evaluation and Reporting of Environmental and Social Performance, pp. 132–150. Greenleaf Publishing (1999)
5. Kunert, A.G.: Environmental Report 1994/95. Kunert AG, Immenstadt (1996)

6. Fava, J., Consoli, E.J., Denison, R., Dickson, K., Mohin, T., Vigon, B. (eds.): Conceptual Framework for Life Cycle Impact Analysis. In: ETAC and SETAC Foundation for Environmental Education, Workshop Report, 1.-7 (1992)
7. O'Riordan, T., Cameron, J., Jordan, A.: Reinterpreting the Precautionary Principle. Cameron May, London (2001)
8. Elkington, J.: Towards the sustainable corporation: win-win-win business strategies for sustainable development. California Management Review, 90–100 (Winter 1994)
9. Gallopín, G.: A systems approach to sustainability and sustainable and sustainable. United Nations Publication, Santiago (March 2003)
10. Drew, O., Shi, Q., Merabti, M.: Policy Translation and Reconciliation Techniques for the System-of Systems Domain, pp. 182–190. Springer, Heidelberg (2009)
11. Bobrow, D.: Artifical Intellignece in Perspective, First MIT Press edn. Elseviers Science Publishers, Amsterdam (1994)
12. Clayton, T., Radcliffe, N.: Sustainability: A Systems Approach. Earthscan Ltd. (April 25, 1996)
13. Rocha, Luis, M.: Complex Systems Modeling (1999), Indiana University Website: http://informatics.indiana.edu/rocha/complex (Retrieved December 13, 2009)
14. Pahl-Wostl, C.: Integrated Assessment III Integrated Modelling (2004/2005), Homepage of Claudia Pahl-Wostl, http://www.usf.uos.de/~pahl/, http://www.usf.uos.de/~pahl/ (Retrieved December 2, 2009)

The Influence of Security Statement, Technical Protection, and Privacy on Satisfaction and Loyalty; A Structural Equation Modeling

Hamid Reza Peikari

Graduate School of Business, University Kabangsaan Malaysia
(National University of Malaysia) Bangi, Malaysia, 43600
omid726@yahoo.com

Abstract. Customer satisfaction and loyalty have been cited as the e-commerce critical success factors and various studies have been conducted to find the antecedent determinants of these concepts in the online transactions. One of the variables suggested by some studies is perceived security. However, these studies have referred to security from a broad general perspective and no attempts have been made to study the specific security related variables. This paper intends to study the influence on security statement and technical protection on satisfaction, loyalty and privacy. The data was collected from 337 respondents and after the reliability and validity tests, path analysis was applied to examine the hypotheses. The results suggest that loyalty is influenced by satisfaction and security statement and no empirical support was found for the influence on technical protection and privacy on loyalty. Moreover, it was found that security statement and technical protection have a positive significant influence on satisfaction while no significant effect was found for privacy. Furthermore, the analysis indicated that security statement have a positive significant influence on technical protection while technical protection was found to have a significant negative impact on perceived privacy.

Keywords: Loyalty, Satisfaction, Security statement, Technical protection, Privacy.

1 Introduction

Customer satisfaction is mentioned as one of the indicators of a company's past, present and future performance[15]. Similarly, in the marketing literature, customers' loyalty is the most important and ultimate outcome [2]. Therefore, exploring the factors determining customers' satisfaction and loyalty is important.

Some studies found common determinants for customers' online satisfaction and loyalty [10, 14 20] indicating that these two concepts are closely associated with each other. Researchers have cited many factors as the antecedents and determinants of customers' online loyalty and satisfaction. For instance, security features of the online stores from which customers make their purchase are among the most cited factors [4, 13]. However, these studies have referred to security attributes of the online store from a general broad perspective and no attempts have been made to explore the direct influence of specific security dimensions on customers' online satisfaction and loyalty.

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 223–231, 2010.

The main objective of this study is empirically investing the direct influence of customers' perceptions about the specific security related factors on their loyalty and satisfaction. This study contributes in the body of knowledge by empirically examining the impacts of new variables on satisfaction and loyalty. The findings improve our knowledge by developing the models available in this field while assist e-retailers and practitioners to formulate more efficient e-commerce strategies.

2 Loyalty

Loyalty is defined as "a customer's favorable attitude toward an e-commerce website that predisposes the customer to repeat buying behavior" [4: 412]. Caceres and Paparoidamis [2] categorized loyalty into behavioral and attitudinal loyalty. According to Caceres and Paparoidamis [2] behavioral loyalty depends on situational factors and conditions while attitudinal loyalty is more important since it is the antecedent of behavioral loyalty.

Satisfaction has been suggested as one of the antecedents of loyalty [4, 8, 10, 11, 14, 15]. It is argued that satisfied customers are more likely to return back to a Website for more purchase [15] implying the positive effect of customers satisfaction on their loyalty. According to the principles of relationship marketing, the more satisfied are customers with a relationship, the more is the likelihood that they are loyal to the relationship [2]. Therefore, it is suggested that:

H1: *Customers' satisfaction has a positive significant influence on their loyalty.*

3 Satisfaction

Gummerus et al. [15] define satisfaction as "a cumulative, attitude-like judgment that is based on customers' past experiences" (p 176). Some authors have stated that there are two dimensions for customer satisfaction: 1- "transaction specific" dimension which refers to the emotional response and satisfaction of customers from their past experience and purchase, and 2- "cumulative customer satisfaction" which refers to a customer's overall satisfaction from dealing with a firm over time [3, 4, 8, 26].

Many determinants of loyalty and satisfaction are common [10, 14, 16, 20, 21]. This implies that the production processes of satisfaction and loyalty are similar. Likewise Website security and privacy attributes have been suggested as the antecedents of satisfaction [4, 24] and loyalty [13]. Therefore, it is hypothesized that:

H2: *Customers' perceived privacy attributes of an online store have a positive significant influence on their satisfaction with the online store.*

H3: *Customers' perceived privacy attributes of an online store have a positive significant influence on their loyalty with the online store.*

4 Privacy

Flavian and Guinaliu [13] define privacy as the ability of individuals to control the terms and condition under which their personal information is stored and used. According to

Chiou et al. [8] privacy is an e-retailer's commitment to protect customers' personal data, online behavior, and credit information. Privacy is one of the antecedents of customers' online trust in a Website [7, 9, 25, 27] and purchase intention [12].

It is expected that the online environment negatively influences customer's satisfaction due to privacy and security concerns [26]. According to Belanger et al. [1], privacy concerns in the online environment include e-retailer's data collection, usage tracking and intentional information sharing with unauthorized third parties. Some online stores ask customers to provide their personal data such as name, address, contact details including email and phone number, and some times more detailed data such as date of birth, passport number, driving license number and so on, which are quite sensitive and make individuals vulnerable in case of any misuse. Therefore, unless customers are confident about the privacy of their personal data, they do not provide any data to the online store and consequently, no purchase in conducted. This indicates the importance of privacy measures in the process of e-commerce diffusion.

5 Security

Security concerns of customers have been cited as one of the barriers of e-commerce diffusion [5, 14, 23]. Security is obtaining customers' data by unauthorized third parties whereas privacy refers to potential misuse of customers' data by merchants [24].

To encourage customers involve a transaction, e-retailers have to demonstrate their ability and willingness to safeguard customers' data and information [20]. Customers' perception and evaluation of the security protection has been categorized into two groups: objective and subjective security issues [6]. Measures such as security policy statement and technical protection refer to customers' objective judgment while overall perceived security refers to their subjective evaluation process [18]. Due to the lack of technical expertise and knowledge, it is difficult for customers to evaluate the Website security measures objectively and tend to evaluate them from the subjective perspective [18]. Therefore, customers' perception about security statement and technical protection mechanism employed by an online store form their overall perceived security measures of the online store. Since some studies have suggested customers' perceived security as the antecedents of their satisfaction and loyalty [4, 13, 24] it is suggested that:

H4: *Customers' perception about the security statement of an online store has a positive significant influence on their satisfaction with the online store.*

H5: *Customers' perception about the security statement of an online store has a positive significant influence on their loyalty with the online store.*

H6: *Customers' perception about the technical protection mechanism of an online store has a positive significant influence on their satisfaction with the online store.*

H7: *Customers' perception about the technical protection mechanism of an online store has a positive significant influence on their loyalty with the online store.*

6 Technical Protection

Technical protection, as the foundation of online transactions, is defined as the technical mechanisms and solutions which aim at protecting customers' security in an online transaction and consequently has a positive significant influence on customers' perceived security and trust [18]. Kim et al. [19] categorized technical protection into security protection and privacy protection. Perceived privacy protection, according to them refers to customers' perception of the likelihood that the e-retailer protects customers' sensitive personal information from any unauthorized access, disclosure and misuse. Such misuses can harm customers in a variety of ways such as spam mail and fraudulent activities. Privacy concerns of customers are positively related to their perceived risk concerning their online transactions with an e-retailer [19]. Perceived security protection however, refers to customers' perception that an e-retailer maintains the authentication, encryption, integrity, and non-repudiation aspects of their data and information [19].

Gummerus et al. [15] believe that there is a close association between security measures of a Website and privacy of its customers. There is a close relation between privacy and security concepts in the minds of customers and they do not distinct them as protection of customers' personal data and privacy depends not only on following a series of ethical practices, but depends also on the reliability and security of the information systems employed by the company [13]. Therefore, it is expected that customers' perceptions about the technical protection solutions of an online store is not only associated with their perceived security, but also is associated with their perceived privacy. Therefore, it can be hypothesized that:

H8: *Customers' perceptions about the technical protection mechanisms of a Website have a positive significant influence on customers' perceived privacy measures of the Website.*

7 Security Statement

Security statement is the information e-retailers provides to customers to ensure the security mechanisms of their Website operations [18]. It is a statement constituting the e-retailers competency and ability to protect the security of customers' sensitive data from the access of unauthorized third parties. It is believed that the presence of security statements in an online store is negatively related to consumers' perceived risk of the online transaction [22].

According to cue utilization theory, when customers visit an online store, they evaluate the merchant's trustworthiness using different cues available on the Website such as security statement. The consistent presence of multiple cues in an online store develops a synergic interaction among them and each of them strengthens the presence of the others, a phenomenon called as cue consistency theory [17]. A convincing influential security statement can ensure that the Website has employed reliable technical protection solutions to protect customers' sensitive data and information. According to cue consistency theory, it is expected that the presence of security statement in a Website influences customers' perceptions of the technical protection mechanism of the Website. Therefore, it is hypothesized that:

H9: *Customers' perceptions about the security statement of an online store have a positive significant influence on their perceived technical protection mechanisms of the Website.*

8 Research Model

To examine customers' perceptions, a quantitative method was employed and a set of structured closed-ended measurement items. As illustrated in Table 1, the scale items were adopted and adapted from published papers. All the question items except the demographic questions were measured on a five point Likert scales.

The data was collected from 337 respondents. They were asked to fill out the questionnaire referring to a familiar e-commerce website from which they have purchased a product or service and no specific Website was mentioned as their reference. It ensures that the respondents have sufficient familiarity with the Website of which they evaluate its features [5].

Table 1. The reliability of the final scale

Variable	Source	Cronbach's Alpha
Loyalty	[11]	0.792
Satisfaction	[11]	0.806
Technical Protection	[18]	0.820
Privacy	[12]	0.839
Security Statement	[18]	0.804

It was found that the majority of the respondents (66.8%) are females. Moreover, it was found that 71.2% of the respondents had the experience of more than one time online purchase. The self reported IT skills of the respondents indicate that 97.3% of them have a self reported IT skills of average or above average level. Both figures of online purchase experience of the respondents and their IT skills indicate that they have a relatively enough knowledge to clearly understand and respond the questions.

9 Reliability and Validity Administration

The reliability of the scale was administered by Cronbach's alpha using SPSS 16 and the scale was revised accordingly. As shown in the Table 2, the revised scale was found highly reliable.

The maximum likelihood estimation technique was applied to examine the validity of the scale using confirmatory factor analysis (CFA). As shown in the Table 2, all the fit indices are above the cut-off levels suggesting an excellent model fit for the measurement model. Convergent validity was examined using the guidelines proposed by Chang and Chen [4] and all the item loadings were found significant (t-values >2.58, p-values <0.001) while all the items loadings were above 0.520. These results ensure the convergent validity of the scale. Moreover, all the paired correlation between the latent variables were found less than 1, indicating the discriminanat validity of the scale. Therefore, the validity of the measurements was ensured.

Table 2. Fit Indices for measurement and structural model

Fit Indices	Recommended value	Result value (Measurement Model)	Result value (Structural Model)
$\chi 2$	NA	243.75	243.976
d.f.	NA	160	161
$\chi 2$/d.f.	≤ 3	1.523	1.515
GFI	≥ 0.9	0.934	0.934
NFI	≥ 0.9	0.917	0.917
NNFI	≥ 0.9	0.964	0.965
CFI	≥ 0.9	0.970	0.970
RMSEA	≤ 0.05	0.039	0.039
RMSR	≤ 0.05	0.033	0.033
AGFI	≥ 0.8	0.914	0.914

10 Hypotheses Testing and Findings

In order to test the hypothesis, structural equation modeling technique was applied using AMOS 5. As illustrated in Table 2, the goodness of fit was examined for the structural model and the results were found above the cut-off levels, indicating an excellent model fit.

To examine the hypotheses, the path analysis technique was used and as shown in Table 3, the results found that loyalty is positively and significantly influenced by customers' perceptions about security statement, and satisfaction but perceived privacy and technical protection have no significant impacts on loyalty. Moreover, the results indicated that customers' satisfaction is significantly affected by their perceptions about security statement and technical protection of the online store, while privacy has no significant influence on their satisfaction. Furthermore, the analysis found significant results for the influence of customers' perceptions about the security statement on their perceived technical protection. However, the results suggest partial

Table 3. Hypotheses and path analysis results

	Hypotheses	**t-value**	**p-value**	**β or γ**	**Results**
H1	Satisfaction → Loyalty	6.377	0.000	0.561	Supported
H2	Privacy → Satisfaction	-0.402	0.687	-0.022	Not Supported
H3	Privacy → Loyalty	-0.622	0.534	-0.331	Not Supported
H4	Security Statement → Satisfaction	2.985	0.003	0.253	Supported
H5	Security Satement → Loyalty	2.843	0.004	0.227	Supported
H6	Technical Protection → Satisfaction	6.166	0.000	0.480	Supported
H7	Technical Protection → Loyalty	0.746	0.456	0.066	Not Supported
H8	Technical Protection → Privacy	-4.242	0.000	-0.271	Partially Supported
H9	Security Statement → Technical Protection	8.168	0.000	0.665	Supported

support for H8, indicating that though technical protection has significant influence on perceived privacy, the effect is negative.

11 Conclusion

This research is the first attempt to study the influence of security statement and technical protection on satisfaction and loyalty and the interrelations between these antecedents since the previous papers had studied the influence of security from a broad general perspective.

The results indicate that customers' privacy has no significant influence on their satisfaction and loyalty. This is inconsistent with the results of some authors [8, 24] who found that security/privacy have significant influence on satisfaction but is consistent with Flavian and Guinaliu [13] who found that privacy/security have no significant effects on loyalty. This implies that privacy is not a direct determinant of loyalty. Moreover, the results suggest that the process of satisfaction production mechanisms and its antecedents are not context free and depend on the context, and respondents' backgrounds, norms and values.

Moreover, it was found that loyalty is a function of customers' perceptions about the security statement and their satisfaction. The positive significant influence of satisfaction on loyalty is consistent with the previous papers [2, 8]. The findings suggest that when customers' observe the security statement in an online store, the feel confident about the commitment of the merchant to maintain the security of their sensitive data from unauthorized access, which reduce their security concerns and perceived risk of disclosure of their data in future transactions with the Website and make them loyal with the Website. A similar phenomenon can be observed in the relation between security statement and technical protection with satisfaction. When customers have a positive perception about the security statement and technical protection mechanism of an online store, they are satisfied with the security services offered by the Website to guarantee their security.

Moreover, the significant influence of security statement on technical protection suggests that when customers observe the commitment of an online store- stated in the security statement- to protect the confidentiality of their data and information, they are ensured about the technical competency and ability of the Website to technically protect their confidentiality against any unauthorized attempts to access their data.

12 Limitations

The sampling method in this study was a convenient sampling and consequently, the findings can not be generalized. Moreover, this paper only studied security statement and technical protection as the specific security related variables and neglected to include more variables such as security seal, authentication, and encryption mechanisms in the model. It is suggested that future research include the above specific variables in the model to study their direct influence on satisfaction and loyalty. Furthermore, the fields of loyalty and satisfaction are not context free and depend on the

respondents' norms and culture. Hence, it is suggested that more studies are conducted in different nations with different cultures, economic situation and e-readiness ranks to compare the perceptions of the nations on this regard.

References

1. Belanger, F., Hiller, J.S., Smith, W.J.: Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. Journal of Strategic Information Systems 11, 245–270 (2002)
2. Caceres, R.C., Paparoidamis, N.G.: Service Quality, Relationship Satisfaction, Trust, Commitment and Business-to-Business Loyalty. European Journal of Marketing 41(7/8), 836–867 (2007)
3. Casaló, L., Flavián, C., Guinalíu, M.: The Role of Perceived Usability, Reputation, Satisfaction and Consumer Familiarity on the Website Loyalty Formation Process. Computers in Human Behavior 24, 325–345 (2008)
4. Chang, H.H., Chen, S.W.: Consumer Perception of Interface Quality, Security, and Loyalty in Electronic Commerce. Information & Management 46, 411–417 (2009)
5. Chang, H.H., Chen, S.W.: The Impact of Online Store Environment Cues on Purchase Intention Trust and Perceived Risk as a Mediator. Online Information Review 32(6), 818–841 (2008)
6. Chellappa, R.K., Pavlou, P.: Perceived Information Security, Financial Liability and Consumer Trust in Electronic Commerce Transactions. Logistics Information Management 15(5), 358–368
7. Chen, Y.-H., Barnes, S.: Initial Trust and Online Buyer Behaviour. Industrial Management & Data Systems 107(1), 21–36 (2007)
8. Chiou, J., Wu, L., Sung, Y.: Buyer Satisfaction and Loyalty Intention in Online Auctions Online Auction Web Site versus Online Auction Seller. Journal of Service Management 20(5), 521–543 (2009)
9. Chiu, C., Chang, C., Cheng, H., Fang, Y.: Determinants of Customer Repurchase Intention in Online Shopping. Online Information Review 33(4), 761–784 (2009)
10. Cristobal, E., Flavian, C., Guinaliu, M.: Perceived E-service Quality (PeSQ) Measurement Validation and Effects on Consumer Satisfaction and Web Site Loyalty. Managing Service Quality 17(3), 317–340 (2007)
11. Dimitriades, Z.S.: Customer Satisfaction, Loyalty and Commitment in Service Organizations; some Evidence from Greece. Management Research News 29(12), 782–800 (2006)
12. Eastlick, M.A., Lotz, S.L., Warrington, P.: Understanding Online B-to-C Relationships: An Integrated Model of Privacy Concerns, Trust, and Commitment. Journal of Business Research 59, 877–886 (2006)
13. Flavian, C., Guinaliu, M.: Consumer Trust, Perceived Security and Privacy Policy Three Basic Elements of Loyalty to a Web Site. Industrial Management & Data Systems 106(5), 601–620 (2006)
14. Flavian, C., Guinaliu, M., Gurrea, R.: The Role Played by Perceived Usability, Satisfaction and Consumer Trust on Website Loyalty. Information & Management 43, 1–14 (2006)
15. Gummerus, J., Liljander, V., Pura, M., Riel, A.V.: Customer Loyalty to Content-basedWeb sites: the Case of an Online Health-care Service. Journal of Services Marketing 18(3), 175–186 (2004)
16. Harris, L.C., Goode, M.M.H.: The Four Levels of Loyalty and the Pivotal Role of Trust: a Study of Online Service Dynamics. Journal of Retailing 80, 139–158 (2004)

17. Hu, X., Wu, G., Wu, Y., Zhang, H.: The Effects of Web Assurance Seals on Consumers' Initial Trust in an Online Vendor: A Functional Perspective. Decision Support Systems 48, 407–418 (2010)
18. Kim, C., Tao, W., Shin, N., Kim, K.: An Empirical Study of Customers' Perceptions of Security and Trust in E-payment Systems. Electronic Commerce Research and Applications 9, 84–95 (2010)
19. Kim, D.J., Ferrin, D.L., Rao, H.R.: A Trust-based Consumer Decision-making Model in Electronic Commerce: The Role of Trust, Perceived Risk, and their Antecedents. Decision Support Systems 44, 544–564 (2008)
20. Kim, J., Jin, B., Swinney, J.L.: The Role of Etail Quality, E-satisfaction and E-trust in Online Loyalty Development Process. Journal of Retailing and Consumer Services 16, 239–247 (2009)
21. Lee, T.M., Jun, J.: Contextual Perceived Value? Investigating the Role of Contextual Marketing for Customer Relationship Management in a Mobile Commerce Context. Business Process Management Journal 13(6), 798–814 (2007)
22. Miyazaki, J., Fernandez, K.: The Antecedents and Consequences of Trust in Onlinepurchase Decisions. Journal of Interactive Marketing 16(2), 47–63 (2000)
23. Peterson, D., Meinert, D., Criswell II, J., Crossland, M.: Consumer Trust: Privacy Policies and Third-party Seals. Journal of Small Business and Enterprise Development 14(4), 654–669 (2007)
24. Ribbink, D., van Riel, A.C.R., Liljander, V., Streukens, S.: Comfort Your Online Customer: Quality, Trust and Loyalty on the Internet. Managing Service Quality 14, 446–456 (2004)
25. Roca, J.C., Garcia, J.J., Vega, J.J.: The Importance of Perceived Trust, Security and Privacy in Online Trading Systems. Information Management & Computer Security 17(2), 96–113 (2009)
26. Shankar, V., Smith, A.K., Rangaswamy, A.: Customer Satisfaction and Loyalty in Online and Offline Environments. International Journal of Research in Marketing 20, 153–175 (2003)
27. Yousafzai, S.Y., Pallister, G.H., Foxall, G.R.: A Proposed Model of E-trust for Electronic Banking. Technovation 23, 847–860 (2003)

"Fiscal Illusion Causes Fiscal Delusion – Please Be Careful!"

Paulo Mourao

Department of Economics, School of Economics and Management,
University of Minho, Gualtar, 4700 Braga, Portugal
{Paulo Mourao,paulom@eeg.uminho.pt}

Abstract. Although fiscal illusion was first articulated by Puviani (1903), it is a current issue, and its consequences are rather serious, as shown in this paper. Fiscal illusion enlarges the public sector but it can, by itself, be a considerable threat to the trust between the State and its citizens, ultimately promoting serious instability in economic cycles and in democratic institutions.

Keywords: Fiscal Illusion, Democratic Institutions, Economic cycles.

1 Introduction

The relationship between the State and the citizens is complex. This sentence may seem a tautology, but this complexity has in fact led to many different conceptions of the relationship between the State and the citizens over history. Recall, for example, the "social contract" of Rousseau, the "dialectic pulsing" of Feuerbach, or the "welfare state" of Hayek.

At the end of the 19th century, the new Italian Republic added to a stimulating debate on the conceptualization of the State. Many intellectuals organized private and public debates on it, classifying the States into three categories (the Monopolistic State, the Paternalist State, and the Voluntary State, as argued by Fasiani, 1941). From these debates and related academic works emerged one of the most interesting schools of economic and political thought in Europe at that time, the Italian School. Formally, they tried to manage the novelty of Paretian principles with a more descriptive side derived from Historicism. The close proximity to the Austrian universities also promoted the insertion of Austrian Subjectivism (mainly Mises and Hayek's points of view) into the articles published in the official journal of this School (*Giornale degli Economisti*).

One of the most original economists of this group was Amilcare Puviani, the "father" of fiscal illusion. His main original contribution was the conception of invisible contracts between the State and its citizens, with fiscal illusion being an institutional product between those who rule and those who are ruled. Because people do not like taxes but accept to be ruled, those who rule need to collect taxes with a minimum of public protest. As Holcombe (2004) notes, the government may be seen as unnecessary for some, but it is inevitable for all. Alternatively, Colbert describes the conundrum as follows: "The art of taxation consists in so plucking a goose as to obtain the largest amount of feathers with the least possible amount of hissing" [Stiglitz, 1999].

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 232–237, 2010.

This text will show that fiscal illusion exists and it damages democracy. Section 2 will introduce the political origins of fiscal illusion. Section 3 will notice how this phenomenon is rather actual. Section 4 will show its main political consequences, namely 'fiscal delusion'. Section 5 will conclude.

2 Origins of a Modern Phenomenon

Fiscal Illusion is the phenomenon that occurs when taxpayers and electors are not aware of the fiscal reality. That is, they do not know how much they receive from the State or how much they pay to it. This is the closest definition to the intentions of Amilcare Puviani, who was first to write about Fiscal Illusion in the early years of the 20th century. In 1903, Amilcare Puviani edited the book *Teoria della Illuzione Finanziaria*. Amilcare Puviani believes that fiscal illusion is not good for the citizens because it enables public expenditures to be obscured, it could be seen as an abuse by the State, and because it deteriorates the bilateral trust between each citizen and his government. Puviani (1903) recognizes that the identification of every financial action carried out by the State was the first step to eradicate Fiscal Illusion:

"Built over centuries, there are very subtle and deep, current coercive systems attentively elaborated by Politics that must be identified by Science" [Fasiani, 1941, pp. 61].

Buchanan (1967, pp. 4) also argues that because taxpayers pay taxes, because voters vote and because rulers are not Plato's philosopher-kings or perfect social planners, Fiscal Illusion exists:

"The omniscient and benevolent despot does not exist, despite the genuine love for him sometimes espoused, and, scientifically, he is not a noble construction. To assume that he does exist, for the purpose of making analysis agreeable, serves to confound the issues and to guarantee frustration for the scientist who seeks to understand and to explain."

More recently, Caplan (2007) suggests that this kind of phenomena can reproduce themselves; they can adopt themselves to national institutions and can become more serious in a silent way.

Formally, Fiscal Illusion can be understood as the result of interactions among rational agents who try to maximize their utilities. However, as O'Brien and Shieh (1990) suggest, the aims of this optimization are achieved with different information restrictions – some agents have more correct information of fiscal values than others do. Depending on the behavior of institutions in political markets, the result of the interactions among these agents tends to be far from the social optimum.

3 A Current Phenomenon

Fiscal Illusion is a current reality. It is rather common that politicians opposing governments use expressions like 'illusions', 'illusionary practices', or 'manipulation over the perception of reality' when they intend to weaken the prestige of incumbents' works.

From a different perspective, diverse polls show that the imperfection of the knowledge of fiscal dimensions is a strong and statistically significant imperfection. The most dated polls are used by Fasiani (1941, pp. 91-92). However, other authors

have also found results that point in this sense using refreshed data (Buchanan, 1967; Caplan, 2007, pp. 57-80). Thus, even in our 21st century, voters do not pay attention to public accounts, making them vulnerable to Fiscal Illusion (as proven by Thaler, 1999, or Sanandaji and Wallace, 2003).

Puviani (1903) classifies Fiscal Illusion as a particular case of Political Illusion. According to Puviani's taxonomy, every time a politician distorts reality, a Political Illusion is generated. It may happen when a politician uses verbal lies, but this is not the only means by which Political Illusions are made. They can also appear when politicians manipulate public reports, show statistics that do not reflect reality, or when they anesthetize the public opinion (Cooper, 2006). Focusing on Fiscal Illusion, it happens every time real fiscal values are distorted. Puviani (1903) and his followers (Fasiani, 1941; Buchanan, 1967; Oates, 1988) detail the many subterfuges used by politicians to generate fiscal illusion, such as resorting more often to indirect taxation than to direct taxes, framing the individual tax burden by numerous moments distributed along the year or postponing the current budget by the increase of public debt.

An attempt has been made by Mourao (2008a) to empirically evaluate the extension of Fiscal Illusion for 68 democracies since 1960. This attempt has produced the Fiscal Illusion Index. The results reveal that Fiscal Illusion varies greatly around the world. Countries such as Mali, Pakistan, Russia, and Sri Lanka have the highest average values over the time considered, whereas Austria, Luxembourg, Netherlands, and New Zealand have the lowest. Regarding the time dimension, between 1980 and 1995, there was a significant decrease in the average value of the index across countries, suggesting a reduction in the adoption of Fiscal Illusion practices during this period. After 1995, the index remained stable in most countries. Mourao (2008a)

As Mourao (2008a) also suggests, this stability is not good at all. It may reflect that fiscal illusion practices are not being reduced but rather that they have been maintained across the democratic countries since 1995.

4 Political and Economic Consequences of Fiscal Illusion

Recently, there is a line of research that has demonstrated that institutions matter for a greater comprehension of economic phenomena (Buchanan, 1967; Alt and Lassen, 2006; or Shi and Svensson, 2006). Therefore, the economic literature has recognized that political business cycles (PBC) are not the same across countries. These cycles differ in length and in amplitude. Driving this heterogeneity are differences in national institutions, such as differences in the maturity of democratic life, in budget transparency and, more recently, in fiscal illusion (Mourao, 2008b). Other serious consequences of fiscal illusion can be observed in economic growth. Mourao (2008c) develops a model that concludes that fiscal illusion tends to harm economic growth in the long run.

Returning to the origins, Puviani (1903) and Fasiani (1941) have already claimed that one of the most serious consequences of fiscal illusion is the abrupt change from a positive view of the State (in the perspective of taxpayers) toward a negative view of the same State. This leads to the degradation of their confidence in public agents and to their growing hostility to the same State. However, I believe that a public revolt in this way, one like a "Boston tea party", resulting from fiscal illusion is unlikely happen in our democratic world for two main points.

First, fiscal illusion is dynamic. I think that people learn from experience but also that incumbents learn from it. If the government reduces direct taxes on my income by one percent and increases the VAT by one percent, I know that I will pay more taxes in the end; in this case, this maneuver does not work for me, and I will not favor the ruling party in the next election for it (however, it may work for those who are not so concerned about public finances). However, if the government introduces a new scheme of taxation, and I do not realize the ultimate extension of it, I can be affected by a more pronounced fiscal illusion the first time I face this scheme. This second time, I begin to doubt the benevolence of the government, and the success of the fiscal illusion produced by this scheme begins to diminish to a point at which the government needs a new scheme to collect more taxes while avoiding my dissatisfaction.

As a result, because fiscal illusion is dynamic, new schemes appear every year. Although electors and taxpayers show increasing resistance to old-fashioned schemes, politicians can introduce new schemes to use the novelty of their "illusions". Then, we can think that fiscal creativity by the State and new taxation schemes are the price set between those who have the power to rule and those who accept to live in democratic institutions. Here, as the risk of hostility to the State increases, so does the fiscal creativity. Note that fiscal reforms happen more often when the popularity or/and the credibility of the Government is low.

Second, some voters and taxpayers also gain from fiscal illusion and thus tend to accept the existence of certain levels of political illusions. This can appear as a paradox, but note that there are taxpayers who are able to receive political rents from the State (for instance, those who are enrolled in fraudulent or corrupt schemes also involving public agents); disregarding these people, it is also interesting that all the others pay (taxes) and do not protest. Additionally, there are those who show a passive role as citizens because of two main reasons: "love for stability" and "aversion to fiscal issues". There are those who like political stability and those who (for idiosyncratic or ethical reasons) do not like loud disturbances (they can change the party or the chief executive through their votes but are not interested in reforming democracy).

"Aversion to fiscal issues" occurs with many people living in democratic countries. People are interested in being employed, being healthy and caring for their families. However, common people are not interested on discussing the ultimate consequences of the VAT, the public debt, the public budget, the size of the government, or the number of public employees (Cooper, 2006). They can protest when waiting for the bus (if they can), but they tend to pay the taxes they ought to pay (according to Cobham, 2005, fiscal evasion is higher in developing countries than in OECD countries). As a result, as the "aversion to doctors" leads some people to avoid medical consultations as long as they (or their health) allow, so does "fiscal issues aversion" leads people to avoid thinking about numbers, budget reports, and public finances (Lipford, 2001). For these people, fiscal illusion is a price for their disinterest in public affairs.

5 Final Remarks towards Fiscal Delusion

Puviani (1903) notes that social scientists have a more powerful role in fighting fiscal illusion than anybody else. He believes that only ethical researchers are able to identify the schemes that produce fiscal illusion. Economists should then denounce them

to show how people are paying more than they receive. As Fasiani (1941) argues, this could be seen as a moral obligation for all those who believe in democratic institutions or in economic progress.

Some other authors have already written in this sense. For instance, Boyle and O'Leary (1997) list a detailed group of suggestions that favor more transparency in public reports and a deep constitutional change in the processes of approving public budgets. Miranda and Picur (2003) or Hendrick, Wu and Jacob (2007) ask for simpler taxation schemes geared toward the perspective of taxpayers. Caplan (2007) believes that better education increases the assertiveness of voting taxpayers.

At last, we can agree with Gustave Le Bon, who says that there is a "portion of hope and illusion without which [men] cannot live" [quoted in Caplan, 2007, pp. 116]. However, we also must ask: "at what point does the illusion become harmful, and who pays for the delusion when the illusionist finishes performing?"

Acknowledgement

The financial support provided by the Portuguese Foundation for Science and Technology under research grant PTCD/ECO/65711/2006 (partially funded by FEDER) is gratefully acknowledged.

References

Alt, J., Lassen, D.: Fiscal Transparency, Political Parties, and Debt In OECD Countries. European Economic Review 50(6), 1403–1439 (2006)

Brender, A., Drazen, A.: Political Budget Cycles in New versus Established Democracies. NBER Working Papers Series 10539 (2004), http://www.econ.umd.edu/~drazen/Data_Sets/Data_Sets.html

Buchanan, J.: Public Finance in Democratic Process: Fiscal Institutions and Individual Choice. University of North Carolina Press, Chapel Hill (1967)

Caplan, B.: The Myth of the Rational Voter. Princeton University Press, Princeton (2007)

Cobham, A.: Tax evasion, tax avoidance, and development finance, Queen Elisabeth House Working Paper No. 129 (2005)

Cooper, H.: Black Voodoo Economics, Harvey Cooper, Lulu.com (2006)

Fasiani, M.: Principii di scienza delle finanze. Giappichelli Editore, Torino, vol. I (1941) (Edition read: Princípios de Ciência de la Hacienda, 1962, Aguilar, Madrid, translator: Gabriel de Usera)

Hendrick, R., Wu, Y., Jacob, B.: Tax Competition among Municipal Governments. Urban Affairs Review 43(2), 221–255 (2007)

Holcombe, R.: Government: Unnecessary but inevitable. The Independent Review 5(4) (2001)

Lipford, J.: How Transparent is the US Budget? The Independent Review 5(4), 575–591 (2001)

Miranda, R., Picur, R.: CFO as Budget Magician: Fiscal Illusion in Public Finance. Government Finance Review, 29 (April 2003)

Mourao, P.: The Consequences of Fiscal Illusion on economic growth. eJournal of Tax Research 6(2), 82–89 (2008c)

Mourao, P.: Towards a Puviani's Fiscal Illusion Index. Hacienda Publica Espanola/Revista de Economia Publica 187(4), 49–86 (2008a)

Mourao, P.: Political Budget Cycles and Fiscal Illusion: a panel data study. In: Marques, H., Soukiazis, E. (eds.) Perspectives on Integration and Globalisation. LIT Verlag, Munster (2008)

Mourao, P.: Quatro Ensaios sobre a Ilusão Fiscal. Ph.D. Thesis. University of Minho, Braga (2009)

O'Brien, J., Shieh, Y.: Utility Functions and Fiscal Illusion from Grants. National Tax Journal 43(2), 201–205 (1990)

Oates, W.: On the Nature and Measurement of Fiscal Illusion: A Survey. In: Brennan, G., Grewal, B., Groenwegen, P. (eds.) Taxation and Fiscal Federalism: Essays in Honour of Russell Mathews, Brennan, G., Grewal, B., Groenwegen, P. (eds.), pp. 65–82. Australian National University Press, Sydney (1988)

Puviani, A.: Teoria della illusione finanziaria. Sandron, Palermo (1903)

Sanandaji, T., Wallace, B.: Fiscal Illusion and Fiscal Obfuscation: An Empirical Study of Tax Perception in Sweden. Master's thesis, Stockholm School of Economics (2003)

Shi, M., Svensson, J.: Political Budget Cycles: Do They Differ across Countries and Why? Journal of Public Economics 90, 1367–1389 (2006)

Stiglitz, J.: On Liberty, the Right to Know, and Public Discourse: The Role of Transparency in Public Life. Oxford Amnesty Lecture, Oxford (1999)

Thaler, R.: Mental Accounting Matters. Journal of Behavioral Decision Making 12, 183–206 (1999)

A Coloured Petri Net Analysis of the Transaction Internet Protocol

Christos K. Georgiadis[1], Ioannis Kokkinidis[1], and Elias Pimenidis[2]

[1] Department of Applied Informatics, University of Macedonia,
Thessaloniki, Greece
{geor,kokkin}@uom.gr
[2] School of Computing, IT and Engineering
University of East London, UK
e.pimenidis@uel.ac.uk

Abstract. The Transaction Internet Protocol (TIP) aims to facilitate e-commerce by enforcing atomicity guarantees in transactions distributed between several autonomous transaction processing systems. In this work, the authors explore a holiday booking scenario in which a customer is protected by the TIP; in such a way as to prevent ending up with a hotel reservation without the requested flight reservation that is enclosed in the same holiday package. TIP defines an approach that makes the commit processing independent of the communication protocol used. There are a number of potential pitfalls that make it useful to provide a formal approach to reason about the behavioral properties of TIP. The authors propose a Colored Petri Net model that allows interactive simulation and verification of correctness properties within the CPN Tools modeling environment. The model can be used in the CPN Tools environment for model checking tasks.

Keywords: E-commerce transaction, CP-nets, Model Checking, Simulation.

1 Introduction

Transaction Internet Protocol Version 3.0 (TIPV3) is a Proposed Standard Protocol [1], specified in detail in [2][3]. It is an Internet standards track protocol for the Internet community, which is envisioned that it will be used mainly for a transaction manager on one Internet node to communicate with a transaction manager on another node. Indeed, in many applications where different nodes cooperate on some work, there is a need to guarantee that the work happens atomically. That is, each node must reach the same conclusion as to whether the work is to be completed, even in the face of failures. TIPV3 presents a simple, easily-implemented protocol for achieving this end: the standard method for achieving atomic commitment is the two-phase commit protocol. A number of two-phase commit protocols have been implemented over the years. However, none of them has become widely used in the Internet, due mainly to their complexity. Most of that complexity comes from the fact that the two-phase commit protocol is bundled together with a specific program-to-program communication protocol, and that protocol lives on top of a very large infrastructure.

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 238–249, 2010.

TIPV3 proposes a very simple two-phase commit protocol. It achieves its simplicity by specifying only how different nodes agree on the outcome of a transaction; it allows (even requires) that the subject matter on which the nodes are agreeing be communicated via other protocols. By doing so, we avoid all of the issues related to application communication semantics and data representation (to name just a few). While it is possible to use this protocol for application programs and/or resource managers to speak to transaction managers, this communication is usually intra-node, and most transaction managers already have more-than-adequate interfaces for the task. TIP uses the TLS transport layer security protocol [4] to restrict access to members-only trusted partners (e.g. to check what remote edges TIP transactions will be accepted and to certify the authenticity of the edge), and to encrypt the TIP commands. While it is not expected that this protocol will replace existing ones, it will be relatively easy for many existing heterogeneous transaction managers to implement this protocol for communication with each other.

In this work the authors explore and analyze a typical TIPV3 transaction, using Colored Petri Nets (CP-nets). The CP-nets have an instinctive, graphical representation that helps to distinguish the basic structure of a complex CP-net model (e.g. how each individual process is interacting). Moreover, each CP-net model includes a network of places, transitions and arcs. All these interact with each other through a set of well defined interconnections, in a similar way as in most programming languages [5][6].

2 CP-Net Diagrams and CPN-Tools

The CP-nets have a typical mathematical representation with a well defined syntax and semantics. This representation constitutes the foundation stone for the determination of the various attributes behavior and the analysis methods. Without the mathematical representation it would have been impossible to produce a steady and powerful CP-net language. However, for the practical use of the CP-nets and their tools, an instinctive comprehension of the syntax and the semantics is all that is needed [5]. CP-nets can be simulated either with interaction or automatically. During the simulation with interaction the user has control over the executed events. It is possible for one to see the results of the individual steps directly in the graphical representation of the CP-net. This means that the user can examine the various states and choose between the possible transitions. The simulation with interaction provides a way to navigate through a CP-net model, examining various scenarios and checking if the model works as expected. This comes in contrast to most of the available simulation packages, which often work as black boxes, where the user defines the input and examines the results, but has very little possibility to understand and evaluate the model based on the way the simulation works [7].

The automatic simulations are identical to program executions. The goal is to make it possible to execute the CP-net model as fast and as accurate as possible, without need for interaction and inspection by humans. However, the user must once again be able to examine the results of the simulation. For this reason, it is often more suitable to use moving graphical representations that provide an abstract and specific picture of the current situation and activity of the system. This is where the importance of CPN-Tools lies: they are tools for processing, simulating and analyzing hierarchical

CP-nets with or without a reference to time. It is a result of research, in the University of Aarhus. CPN-tools have strong capabilities in simulating and analyzing discrete event systems [8]. They combine a powerful functionality with a very friendly user interface, and contain improved techniques of interaction, as well as a different type of graphical representation that keeps the user informed for the status of syntactical checking and simulations [9]. CPN-Tools for simulating the TIP protocol have been used. The aim is to simulate a representative scenario (case study) of this protocol, in order to analyze its behavior.

3 Analyzing an Indicative TIPV3 Transaction

The TIPV3 transaction that is to be presented includes the following four participants:

- Customer (C),
- Agency (A),
- Services of Airlines (M1) and
- Services of Hotels (M2).

The conjunctive link between the customer and the various services of airlines and hotels is the agency (A), which in the substance executes debts of the Trusted Third Party in the transaction and ensures the integrity, the reliability and the safety of the transaction. The basic steps of the message exchange protocol are as follows:

Connection process

{1} C = > A	Connection Request.
{2} A = > C	Response.
{4} A = > M1	Connection Request.
{5} M1 = > A	Response.
{6} A = > M2	Connection Request.
{7} M2 = > a	Response.

Correlation of transaction with the connection

{3} C = > A	Reservation of flight and accommodation.
{8} A = > M1	Flight reservation request.
{9} A = > M2	Accommodation reservation request.
{10} M1 = > A	Completion of flight reservation.
{11} M2 = > A	Completion of hotel reservation.
{12} A = > C	Available reservation "packets".
{13} C = > A	Booking commitment and payment.
{14} A = > C	Dispatch of the e-Ticket. Recording of the transaction.

The analysis of the above steps of protocol's message exchange has as follows:

{1}, {2}: the primary member, in our case the customer, and the secondary member, the agency, agree in some version of protocol. Steps {4} and {5} for the agency and the Airline Company and steps {6}, {7} for the agency and the hotel work in a similar way. These steps are executed after step {3}. No reference has yet been made in the transaction.

{3}: the customer dispatches a PUSH command to the agency asking it to execute the process of reservation. The connection henceforth has been related with a transaction and

it can be completed either with a one phase commit protocol or with a two phase commit protocol. The same process is followed by the agency in steps {8} and {9} with the airlines and the hotels respectively, asking from the companies to check the availability of the reservation. In steps {10} and {11} registration and commitment of the reservation of the air ticket and the room of the hotel is being made from the corresponding services. This work of the air companies and hotels can be executed simultaneously. In steps {12}, {13} and {14} the customer asks either for the rejection or the commitment of the transaction from the agent.

The communication in the TIP is based on a protocol of command-response. The primary member sends a command and the secondary can select an answer that is related with the command. In the particular example, the TIP command that is used for the connection by the members is IDENTIFY, while the commands that are related with the transaction are ABORT, COMMIT, ERROR, PREPARE, PUSH and RECONNECT (the analysis of the responses of the particular commands are as defined by the TIP protocol specification - RFC). Connection commands MULTIPLEX and TLS, and transaction command QUERY are omitted for reasons of simplification of the TIP_CPnet, while the transaction command BEGIN is omitted because it can be presented as a COMMIT command after a PUSH command in the state ENLISTED. This COMMIT command presents a one phase commit protocol in contrast to the COMMIT in the PREPARED state, where a two phase commit protocol is presented.

The protocol determines a variety of situations in which one node can result. For each situation a subset of commands of the protocol are allowed in order to change the state of the transaction. Fig. 1 presents the state space of the connection. It is important to comprehend that these states concern the connection and not the transactions.

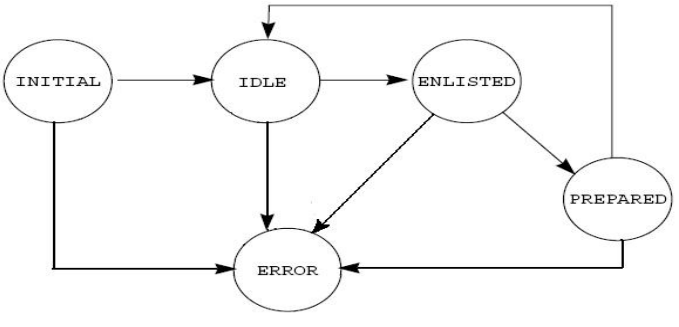


Fig. 1. Valid States of the TIP Connections

A diagram of a successful transaction using the PUSH model appears in fig. 2. It also displays the commands and the responses in order to synchronize the state in both nodes. All commands that are invalid in the current state lead a transition to the ERROR state. This state can also occur from the dispatch of an ERROR command.

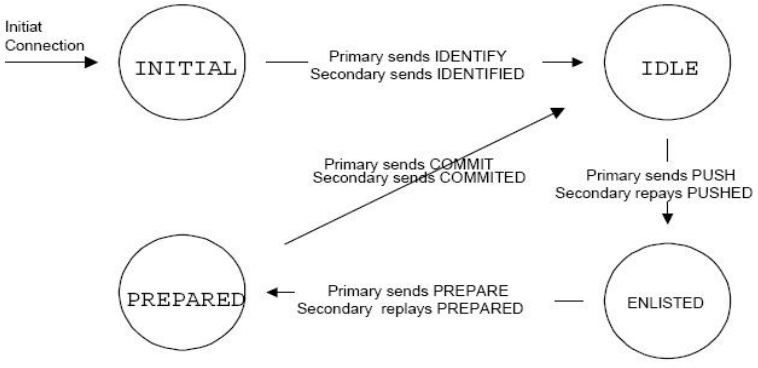


Fig. 2. A Diagram of the States of a Successful (committed) TIP Transaction

3.1 The Structure of the Model

Analyzing our proposed model, we may distinguish the following categories for places that are to be used:

- Places that represent the contracting members.
- Places that represent the one phase commit protocol.
- Places that represent the two phase commit protocol.
- Places that represent the result of transaction.
- Places that represent the channels of communication.

Fig. 3 displays all types of data that are used by the model that is presented. Colset State represents the various states in which a contracting member can be found. All of the members initially begin from the state INITIAL. The state of the secondary changes with the arrival of a command, while the state of the primary changes with the arrival of a response. Colset Command represents the commands that can be dispatched by the primary member and the likely responses that can be received from the secondary. Each command should be given in the right state.

Colset Members defines the three members that are connected with the agency. Customer (C), Airline companies (MA) and Hotels (MH). Colset AState implies the state of connection of the contracting members from the side of the agency. Similarly Colset ACommand represents the command that was received by the agency and by which member. Finally Colset APC indicates if the protocol of transaction being used by the agency and the companies is a one phase commit protocol or a two phase commit protocol.

The variables pc1, pc2, wait and tf are used for the determination of the 1pc or the 2pc protocol. Variables st and st1 imply the state of members, while variables cmd and M imply the commands and the member respectively. Colset Ten0 and Ten1 with their corresponding variables s and t, as well as the function fun OK(s, t) are used clearly for reasons of demonstration of error states in the reception of commands, in order to change the development and flow of the CP-net.

```
Declarations
  Standard declarations
    colset E = with e;
    colset INT = int;
    colset BOOL = bool;
    colset State = with Initial | Idle |
        Error | Enlisted | Prepared;
    colset Command = with ABORT | ABORTED | ERROR |
        COMMIT | COMMITED | IDENTIFY | IDENTIFIED |
        PUSH | PUSHED | NOTPUSHED | ALREADYPUSHED |
        RECONNECT | RECONNECTED | NOTRECONNECTED |
        PREPARE | PREPARED;
    colset Members = with C | MA | MH;
    colset AState = product Members * State;
    colset ACommand = product Members * Command;
    colset APC = product Members * BOOL;
    var pc1, pc2, wait, tf : BOOL;
    var st, st1: State;
    var cmd: Command;
    var M: Members;
    colset Ten0 = int with 0..100;
    colset Ten1 = int with 1..100;
    var s : Ten0;
    var t : Ten1;
    fun OK(s: Ten0, t : Ten1) = (t <= s)
```

Fig. 3. The Variables of the Model

3.2 Description of the TopLevel CP-Net

In the following fig. 4, the TopLevel level CP-net is presented, which has a hierarchical structure. The schema consists of five transitions that represent the customer, the individual companies and the intermediary agency which has a FrontEnd for the management of transactions with the customers and a BackEnd for the management of transactions with the companies. Each one of these transitions is analyzed in an individual lower level CP-net and will be developed more analytically in a following section.

The schema also distinguishes a set of frames, some of which represent the channels of communication of the model such as the CustomerToAgent, AgentToCustomer, AgentToMerchant and MerchantToAgent. Their names also determine the direction of the data flow between the transitions. The data that they transport are TIP commands and their responses. There are also places such as Customer, Agent, Airlines and Hotels which are used in order to display initially the state in which the connection is found and the transaction later on. Places ReservationMade, ReservationAborted, TransactionError and CustomerError use Boolean variables and imply the evolution of the transaction. The place DumpTransaction also uses Boolean variables and when it is true the transaction is aborted and all the communication channels are cleared. Finally the place TransactionResult is of type Command and presents the final situation of the transaction, COMMITED, ABORTED and ERROR.

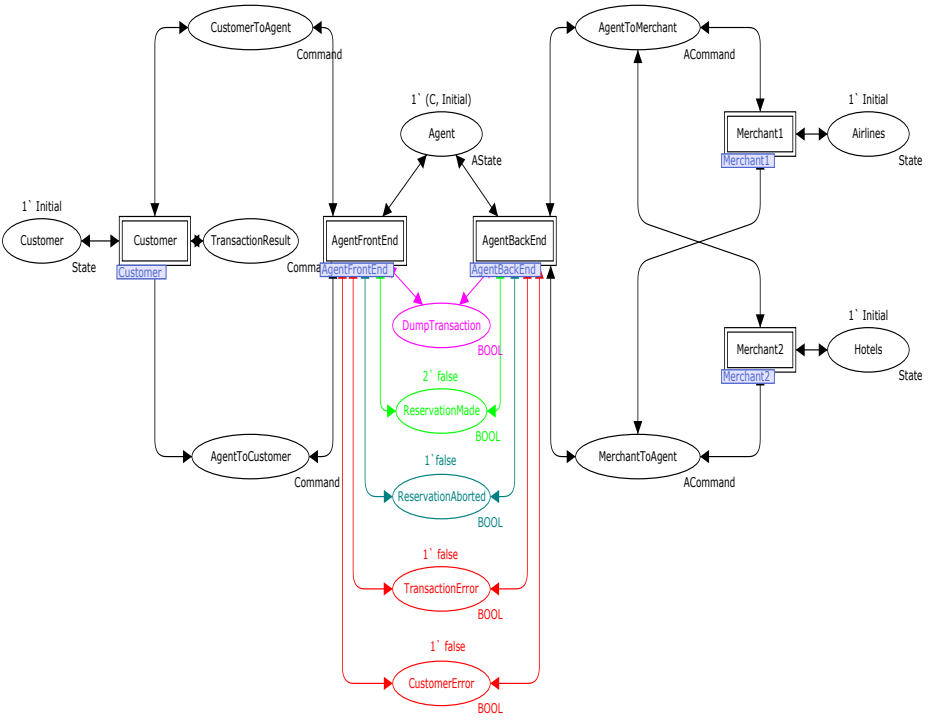


Fig. 4. The TopLevel Level of the Hierarchical Model

3.3 Description of the Customer Model

The customer model (in fig. 5) is consisted of eleven places. The place Customer initially declares the state of the connection and then the state of the transaction. Its initial state is Initial. The places CustomerToAgent and AgentToCustomer present the communication channels between the customer and the agency. All the commands are sent through these channels. Places C1PC and C2PCS of Boolean type, display if the transaction is incarnated with a one phase commit protocol, or if it is incarnated with the use of a two phase commit protocol. Place CWAIT prevents the customers' TM from resending the same command or any other one until it receives answer from the agency. Places CSR1 and CSR2 are used in order to simulate the case at which a command is lost before it reaches its destination. Places CSR3 and CSR4 are used as variables that determine the Error rate that can occur either at the connection or at the inspection of an incoming command. They also determine how often can a transaction be Aborted: either because the Customer does not want to commit the reservation, or because he didn't pay the correct amount, or even because no available service exists. Finally the place TransactionResult is the place in which the final result of transaction is presented.

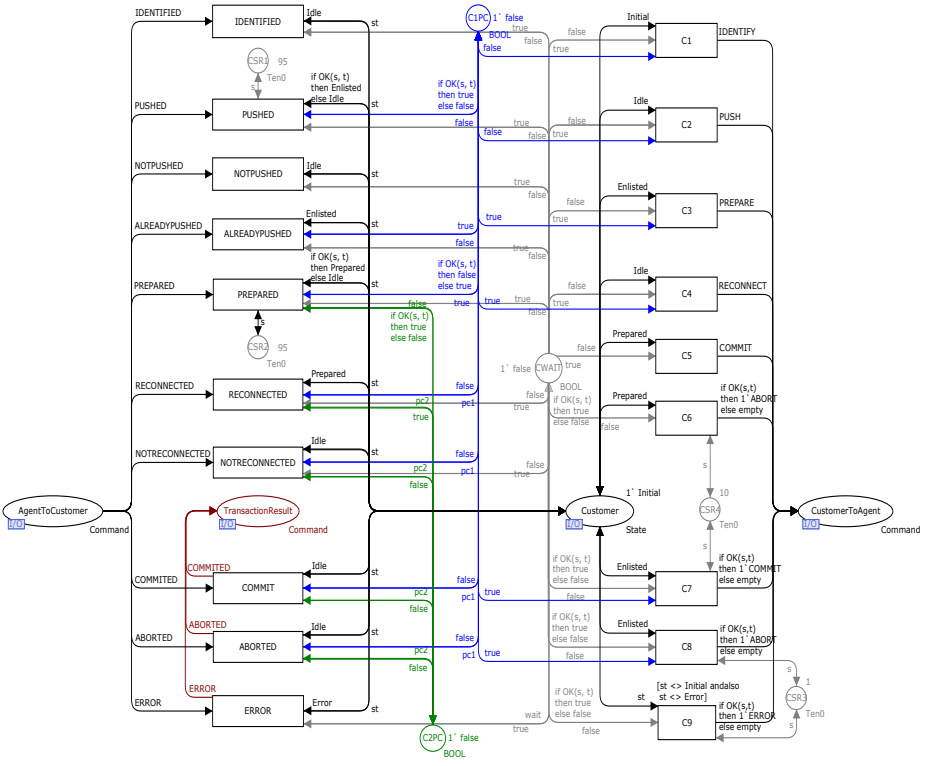


Fig. 5. The Customer Model

In the model there are also nine transitions, named C1 to C9, the purpose of which is to dispatch the suitable commands to the agency, depending initially on the state in which the connection of customer with the agent is and later on the state in which the transaction is. On the other hand, in the input channel, the transitions translate the commands "responses" which are sent by the agency and respectively change the state of the customer. It should be noted for one more time that the state of the customer (primary) changes as soon as it receives a response, in contrast to the agency (secondary) which changes state with the arrival of a command.

More analytically, beginning from the state INITIAL, the customer dispatches an IDENTIFY command via the first transition C1. The common version of the protocol is determined and an IDENTIFIED command is received from the agency. If there is no common version of the protocol the connection is lead to the state of failure (ERROR state). The customer passes in the IDLE state. In this state he can dispatch a PUSH command (C2 transition) seeking from secondary to do some work of the type reservation. Likely responses to this command are PUSHED, all have been undertaken and the process of reservation has begun. Command NOTPUSHED as a response implies that it was impossible to begin the process of reservation. A new PUSH command must be dispatched. The ALREADYPUSHED command arrives at the customer, if one of the two previous responses is lost and command PUSH is

again dispatched. After the arrival of such a response we leave the description of the state of the connection and pass to the description of the state of the transaction. The current state of the customer is ENLISTED.

In this state the customer can select to use a one phase commit protocol and dispatch a COMMIT or an ABORT for committing or aborting the transaction respectively. In the ENLISTED state, the customer can dispatch a PREPARE command in order to complete the transaction using a two phase commit protocol. A likely response to this command is PREPARED. If for any reason this response is lost, the customer moves to the IDLE state and will have to dispatch a RECONNECT command in order to reconnect. The responses that can be received are RECONNECTED, therefore the customer passes henceforth in the PREPARED state and is ready to ask the agency to commit the transaction, and NOTRECONNEDTED where the secondary doesn't know any more about the transaction and thus both of the members move to the IDLE state. Finally the customer dispatches the COMMIT command or the ABORT command depending on whether he wants to commit the transaction or abort it.

3.4 Description of the AgentFrontEnd Model

The AgentFrontEnd model of the agency includes ten places. It represents the model which "listens" for reservation requests from potential customers. The place Agent defines the state in which the agency is found, with regard to its connection and transaction with the customer. The same place is used in the next paragraph for the same purpose, for its connection with the air companies and the hotels. The difference between this place and the first place concerning the customer is that the particular place keeps information both for the state of the connection and for the member that owns this state. Its initial state is INITIAL and listens for the potential customer.

The places CustomerToAgent and AgentToCustomer represent as in the previous paragraph the channels of communication between the agency and the customer. Place ReservationMade becomes true when the agency receives a COMMIT command and responds with the COMMITED command. In order for this to happen, it is supposed that the airline companies and the hotels have both answered previously with COMMITED in the reservation request that the agency has sent separately to each member. If for any reason, one of the two members (airline companies or hotels) respond with ABORTED then this is enough in order for the transaction to be aborted in the whole. If by any chance a command is not recognized or has not been published in a suitable state, or if any other problem with the connection of either contracting member exists, then the place TransactionError or the place CustomerError becomes true and the state of the connection becomes ERROR.

The place DumpTransaction becomes true when the transaction is aborted or when the agency does not know about the transaction any more. In other words, when it is asked to respond to a RECONNECT command and responds with a NOTRECONNECTED. Places SR3 and SR4 are used for the same reasons that were reported in the customer model. The model also includes fourteen transitions, where each one receives a response, changes the state of the connection and the state of the transaction of the agency and dispatches a respond to the customer.

3.5 Description of the AgentBackEnd Model

The model of the agency AgentBackEnd resembles that of the customer with minimal differences, since its way of operation is similar. It includes fourteen places and nineteen transitions. The transitions do not differ from those of customer. Differences are in places ReservationMade, ReservationAborted, TransactionError and DumpTransaction, which were analyzed in the previous section. The places AgentToMerchant and MerchantToAgent represent the channels of communication between the agency and the various services. As it appears in the model structure in the next picture these channels are common for all the involved services and the commands are undertaken based on the member to which they are dispatched. Messages sent through these channels are of type Member*Command.

Places ASR3 and ASR4 are used for the same reasons that were mentioned in the customer model. Place A1PCAllowed could be completely removed including the command places ABORT and COMMIT that are connected to it. Its role is to manually select if we want the model to use the one phase commit protocol. The initial value of the place does not allow the model to use this protocol. All the transactions between the members, agency and services, are set to use the two phase commit protocol. In an opposite case we can set the value of the place True.

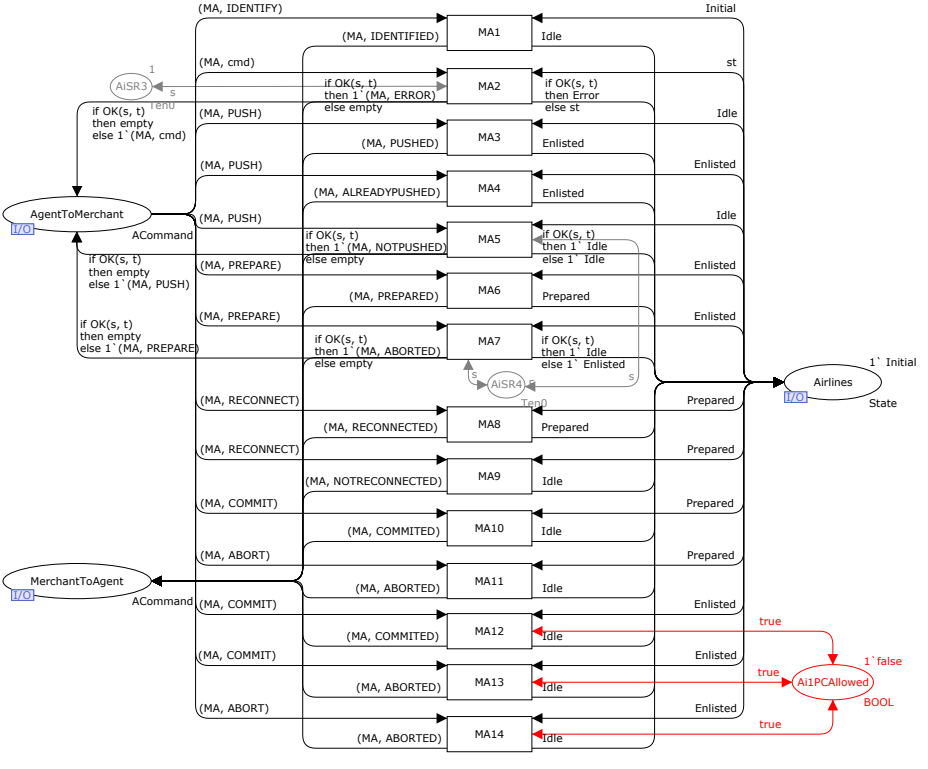


Fig. 6. The Airline Company (Merchant1) Model

3.6 Description of the Airline Company (Merchant1) Model

The airlines model is constituted by four places. The place Airlines determines the current state of connection and transaction of the service with the agency. Places AgentToMerchant and MerchantToAgent represent the channels of communication with the agency. The message that enters the input channel is of type Member*Command. This is how the message is distributed with regard to the destination of the command. The transitions that receive the commands change the current state of the connection or transaction, in a similar way to the one that we described in the AgentFrontEnd model.

Places AiSR3 and AiSR4 are used for the same reasons that were mentioned in the customer model. This model contains also the place Ai1PCAllowed, which deters the model from using the one phase commit protocol. The Hotel Services (Merchant2) model is similar to the previous one: the only difference is that the names of the places AiSR3 and AiSR4 have been changed to MSR3 and MSR4. These places are also used for the same reasons that were mentioned in the customer model.

4 Conclusion

In this work, a holiday booking scenario based on the use of the TIP protocol that enforces atomicity guarantees in e-commerce transactions has been studied. In doing so the authors used a Colored Petri Net model that provides a formal analysis alternative for verifying transaction correctness properties and understanding the behavioral properties of the TIP by interactive simulation.

Using the CPN-Tools, it is possible to perform a set of predefined queries that analyze the structural properties of the model's state space. The first part of the generated report in our TIPV3 case study, presents certain statistical information that concerns the state space graph. The generated reachability graph includes 6136 nodes and 670265 arcs which represent all the different states in which this particular model can be found. There is also information for the corresponding graph with the strongly connected nodes, termed as Scc Graph, which includes 5500 nodes and 63925 arcs. Moreover, the second part contains the Home Properties, which are capable to provide information about markings or sets of markings to which it is always possible to return: it appears that our model does not contain any such markings, as report's Home Markings is 'None'.

Finally, at the third part of the report, the Liveness Properties provide information regarding the number of Dead Transition Instances, Live Transition Instances, and Dead Markings. The Dead Transition Instances display the number of "dead" transitions, in other words transitions that are not executed for at least once in any state of the model. The Live Transition Instances are transitions that are continuously active. No such instances on both of these categories are found in our model's CPN-Tool report. On the other side, 482 Dead Markings were found (nodes without an executable transition). But this might mean either, that they are final states of the protocol or deadlocks and consequently they must be further examined. The authors are currently working on a new extended CP-net, to incorporate new features and improvements.

References

1. Reynolds, J., Ginoza, S. (eds.): Internet Official Protocol Standards, Network Working Group RFC 3700, Standards Track, Internet Engineering Task Force (2004)
2. Evans, K., Klein, J., Lyon, J.: Transaction Internet Protocol – Requirements and Supplemental Information, Network Working Group RFC 2372, Standards Track. The Internet Society (1998), http://www.faqs.org/rfcs/rfc2372.html
3. Lyon, J., Evans, K., Klein, J.: Transaction Internet Protocol Version 3.0, Network Working Group RFC 2371, Standards Track. The Internet Society (1998), http://www.faqs.org/rfcs/rfc2371.html
4. Katsaros, P.: A roadmap to electronic payment transaction guarantees and a Colored Petri Net model checking approach. Information and Software Technology 51(2), 235–257 (2009), http://dx.doi.org/10.1016/j.infsof.2008.01.005
5. Jensen, K.: Coloured Petri Nets: Basic Concepts, Analysis Methods and Practical Use. Monographs in Theoretical Computer Science, vols. 1-3. Springer, Heidelberg (1997)
6. Kristensen, L.M., Christensen, S., Jensen, K.: The Practitioner's Guide to Coloured Petri Nets, CPN Group, Department of Computer Science, University of Aarhus, Denmark. Springer, Heidelberg (1998)
7. Katsaros, P., Odontidis, V., Gousidou-Koutita, M.: Colored Petri Net Based Model Checking and Failure Analysis for E-commerce Protocols. In: Proceedings of the Sixth Workshop and Tutorial on Practical Use of Coloured Petri Nets and the CPN Tools (CPN 2005), DAIMI PB-576, University of Aarhus, Denmark, pp. 267–283 (2005)
8. Georgiadis, C.K., Baltatzis, D., Pangalos, G.: Secure mobile agent environments: modelling role assignments. Int. J. Electronic Security and Digital Forensics 1(3), 249–267 (2008)
9. CPN2000 Project: Coloured Petri-Nets Tools (2000), http://www.daimi.au.dk/CPnets/CPN2000

Identification of the Required Security Practices during e-Government Maturity

Ali Shayan, Behnam Abdi, and Malihe Qeisari

Department of Information Technology Management, Humanities Faculty,
Tarbiat Modares University, Iran
{Ashayan,Babdi,M.qeisari}@modares.ac.ir

Abstract. In spite of the e-government benefits, there are some problems during its successful implementation. One of which is information security issues. In this paper, attempts will be made to illustrate the main practices of information security management in each stage of e-government maturity. This study is based on Delphi technique which carried out in two rounds. Based on the literature, a questionnaire was developed and distributed among 38 experts in the first round. In the second round, 12 experts participated. The IQR (Interquartile ranges) was calculated and it founds that the consensus is convenient. According to the results, trends can be depicted in the security practices which have implications for security vision, policies and practices during e-government maturity. The findings suggest that dealing with few aspects is not sufficient, and comprehensive integrated system of information security management should be regarded according to the specific circumstances of the organizations.

Keywords: E-government Maturity, Information security management, Security policy and practice, Delphi technique.

1 Introduction

During the last decade a revolution in ICT is being witnessed, which changes the daily life of people [1]. these changes being transformed into new forms of government, namely, e-government. E-Government generally refers to the use of information and communications technology (ICT) to change the structures and processes of government organizations [2,3]. E-government practices can be examined in terms of the interactions between sectors of government, business, and citizens. Many government agencies disseminate data to interested parties [4, 5]. The importance of e-government has been widely recognized [6]. Some regard e-government as a powerful tool for improving the internal efficiency of the government and the quality of service delivery as well as enhancing public participation [7]. It should enable better outcomes, quality services, and greater engagement with citizens [8].

In spite of the benefits of developing E.G, there are some problems during and after its successful implementation. Lack of attention to information security causes decline of trust and confidence among citizens towards E.G. Organizations attempt to secure their own IT environment, but they have little control over the IT systems with

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 250–262, 2010.

which they link. In this paper we outline and illustrate the main practices of information security management in each stage of e-government maturity according to the Layne and Lee's model [9] and its extension done by Anderson and Henriksen [10].

2 The literature

In this part, we review the main concepts of e-government and its security challenge as well as the maturity model that which is used in this article. Furthermore, the information security issues, as one of the key problems of today's organizations and its importance in the e-government implementation, are discussed. At the end of this section, we review the previous attempts for identifying the required security practices for e-government and their drawbacks.

2.1 E-Government and Security Challenge

There is no single, widely agreed upon definition of electronic government. Academics have suggested various definitions for e-government. Simply speaking, e-government means the communication between the government and its citizens via computers and a Web-enabled presence. Whitson and Davis have defined e-government as implementing cost-effective models for citizens, industry, federal employees, and other stakeholders to conduct business transactions online [11]. It is also an efficient way of conducting business transactions with citizens and businesses and within the governments themselves. Electronic government comprises the use of modern ICTs to deliver public services to citizens and businesses [12]. A successful implementation of E-Government requires careful planning in many regards. The integration between the new E-Government information systems and the existing internal systems has to be redefined in terms of IT elements and business processes. Successful E-Government initiatives become possible only when managers in government agencies realize and deal with these challenges in an effective manner [13].

The advent of e-government creates both opportunities and challenges for government agencies. E-government projects face multiple and complex challenges [14]. Some of the most important problems of successful developing e-government are:

- Security: Security issues in e-government can include protecting against hackers and viruses, ensuring integrity of electronic records, halting authorized sharing or disclosure of information, and preventing the interception of information.
- Privacy: Many e-government systems collect, store, and use personally identifiable information about users of the services or even visitors to the Websites, therefore the concept of privacy for the individual is notable.

Protecting personal privacy and implementing appropriate security controls are reported by U.S. General Accounting Office as challenges to implementing e-government as well [15].

2.2 E-Govenemnt Maturity Models

There are several models about e-government maturity. In 2001, layne and lee suggested a four stages model for developing fully functional E-government. These stages are (1) cataloguing, (2) transaction, (3) vertical integration, and (4) horizontal integration [9]. But because of the government's new services and appearance of new technologies, Anderson and Henriksen have extended the pervious model. Their model has four stages includes: (1) cultivation, (2) extension, (3) maturity, (4) revolution [3]. The cultivation phase is analogues the vertical and horizontal phases, so this study considered seven stages for e-government maturity as follow:

Stage 1-Cataloguing: initial efforts of state governments are focused on establishing an online presence for the government and information. Many state governments' efforts on Web development -state Website- and forms-online initiatives belong to this stage. Governments do not have enough Internet expertise, and prefer to minimize the risk by doing a small project. Also, Citizens would still use existing service processes such as a phone call, in-person standing in line, etc, but to a lesser extent.

Stage 2-Transaction: In the second stage, e-government initiatives focus on connecting the internal government system to online interfaces and allowing citizens to transact with government electronically. At this stage, e-government efforts consist of putting live database links to online interfaces. As the quantity of these e-transactions increase, governments will be pressed to integrate the states' systems with these Web interfaces, or build online interfaces directly connected to their functional intranet. This stage is the beginning of the e-government as a revolutionary entity changing the way people interact with their government.

Stage 3-Vertical integration: At this stage, the focus is now moving toward transformation of government services, rather than automating and digitizing existing processes. Central and local counterpart systems are expected to connect or, at least, communicate to each other. Physically, this may be integrated as a central database or a connected Web of databases communicating with each other.

Stage 4- Horizontal integration: In this stage, databases across different functional areas will communicate with each other and ideally share information, therefore information obtained by one agency will propagate through out all functions. In addition, citizens could conduct business across a wide variety of requirements. Also, a real one stop shopping for citizens prepared and using Intranet within government become important.

Stage 5-Extension: This stage is the extension usage of intranet and adoption of personalized Web user interface for customer processes. the Web user interface is targeted towards the end-users rather than other public authorities or the agencies themselves. The ambition of having a user interface for the end-users shines through the actual Web site. While this is a key difference between stages 4 and 5, this ambition also presents a key failure risk and precipitates costly user interfaces, no integration with other systems, expensive maintenance, and fading out of old software and data format.

Stage 6-Maturity: Stage 6 is the stage where the organization matures and abandons the use of the intranet, have transparent processes, and offers personalized Web interface for processing of customer requests. The Internet and intranet have merged and the key concern is to use IT to lower the marginal costs. The homepage is feeding

information from other institutions to the users online. The Web site is organized to solve problems and requests rather than presenting formal organizational structures and general information. On-line self-service is a key priority in this phase.

Stage 7-Revolution: this Stage characterized by data mobility across organizations, application mobility across vendors, and ownership to data transferred to customers. The employees' actions can be traced through the Internet and there is information available online about progress. This is possible through intra and extra organizational mobility of data and services the ambition is to transfer data ownership and the orientation of data base infrastructure to the end-users.

2.3 Security

Being secure is composed of cluster of positive (safety, confidence) or negative (freedom from anxiety or fear) affects. [16]. information security is and has always been the discipline to mitigate risks impacting on the confidentiality, integrity and availability of a company's IT resources, [17]. Information security has been defined as encompassing systems and procedures designed to protect an organization's information assets from disclosure to any person or entity not authorized to have access to that information [18]. Information security management is concerned with ensuring business continuity and minimizing business damage by preventing and minimizing the impact of security incidents that threaten an organization's information assets [19]. Entities such as hackers, terrorists, disgruntled employees and business competitors are on the lookout for any vulnerability and may seek to exploit found weaknesses for psychological, political or economic advantage [20]. Privacy and related security issues must be adequately addressed in government IT initiatives [21]. it highlights the range of highly complex and diverse challenges public managers must face as they work in the e-government arena. Findings indicate that companies are not proactively tackling information security management [22]. Requirements for confidentiality, integrity, availability, authenticity, and access control have become more differentiated, but the ability to meet these requirements apparently has not kept up. One of the most problems commonly faced by such security officers is lack of authoritative source of guidance [23]. The aim of this article is preparing such guidance for public sector managers that they should draw on considering their organization situation through determining practical information security requirements.

2.4 Security in e-Government

The information security should be considered in delivering services to the citizens, especially information sharing. Government officials can transmit information in real time or near real time, provide notification to numerous offices simultaneously, compare or merge databases, and integrate seemingly disparate pieces of information [24]. States include a great variety of information policies under the general umbrella of security: configuration management, network security, physical security, security training, personnel security, digital certificates and public key infrastructure (PKI), information security, enterprise security, and authentication are included [25]. The price we have to pay is the complexity introduced in the design of the security mechanisms required for protecting several heterogeneous information systems and ensuring user privacy

[26,27]. A high level of confidence and trust among all users (citizens, businesses and government) will be the foundation of a successful e-government initiative. So, a new framework for identifying and organizing the security requirements are required which can facilitate the development of a unified e-government security policy. Lambrinoudakis et al. in 2003 identified security requirements for an integrated e-government platform. They chose some dimensions of e-government such as e-voting and e-university, and they assessed information security risk and their security requirements [26]. As they mentioned themselves, it is clarified that neither services cover the entire spectrum of e-government services, nor the entire lists of security requirements have been utilized here. But the main shortcoming of their work is that they took this supposition that the e-governments is in a fix state of maturity, whereas each government is in a stage of developing their services in electronic manner. As a result, this study tries to determine information security requirements for each stage in e-government maturity as a basis for policy making.

3 Methodology

In this section, we describe the methodology for identification of the required security practices during e-government maturity. First, the literature about the security practices during e-government maturity is reviewed. An initial list of the most important security practices was prepared and five experts selected the most important practices between them. Then, for the first stage, we offered eighteen practices to all experts. After the first stage of the research, according to the expert's suggestions, modifications have been made. Consequently, the final version of the questionnaire comprised of twenty factors. Table 1 includes the twenty practices and their definitions.

Table 1. Practices and their definitions

Practice	Definition
Asset management	It includes recognition of organization information resources and prioritizes their security importance
Information security risk management	Calculation of threats, vulnerabilities and their possibilities. It considers modification, destruction, Disclosure, interception, interruption and fabrication. Then, it evaluates solutions
Business Adaptation of information security	Information security adaptation according to the business requirements and strategies, cost-benefit analysis and collaboration with other departments in the organization
information security policies and strategies	Determining strategies and annual planning for information security to have a similar perspective in the whole of the organization
Incident management	To ensure continuity of business in the event of major failures or disasters.

Table 1. (*Contionued*)

Environment scanning	Investigation of environment and predicting of the future organization situation for security readiness
implementing information security management standard	Selecting and implementing information security management standard to have tools for measuring current security situation of organization
Secure information system development	Paying attention to the security criteria and standards in the information systems development
Information encryption	Codifying information when it is necessary for preventing unauthorized person's access
Database security	Database information protection while it's being used by the users
Network security	Information protection during its transportation in the networks such as internet and intranets
Information system security	Observing security concerns during the implementation o information systems and applications
access level of Information	Determining customers' and employees' roles in relation with information resources. These roles include reading, adding, removing or modifying information
Physical information security control	Preparing the location of information safety and preventing physical threats
Legal control of e-presence	Adjusting each security practices in the organizations to the national and international rules and protocols
Audit management	Controlling the practices related to security relevant practices for information certification and accreditation
Knowledge management of information security	Security Knowledge recognition, acquisition, documentation, sharing and updating between employees and customers.
Information security Organization	Determining chief security officer and top managers, preparing security procedures and guidelines, organizing processes and human resources.
Human resource training	Improving human resources' awareness of security issues
Security culture establishment	Establishing security culture and atmosphere among staff and costumers for obeying organization security rules

Then, following the Delphi technique, a panel of experts was utilized to determine the importance of the 18 practices in each seven phases of e-government maturity model according to the Likert's continuum (that rank 1 refers to the least importance and 5 refers to the most importance). The 18 practices were followed by a blank space so that experts' opinions regarding other required practices could be added, as well. The most important step of Delphi technique is selecting the respondents [28]. Sackman argues that, a panel of experts composed by people with similar backgrounds and interests may tend to comprise elite with a vested interest in promoting the area under Delphi investigation [29]. To do so, attempts were made not to focus on the knowledge and expertise

of theorists merely, but also the experience of practitioners was regarded as a determining factor in selecting the experts. The experts were selected from two different areas: academic and business. As a result, the findings can be useful for exploitation in both these areas. Fifty percent of the experts had academic experience in the superior universities of Iran, and the others were from business fields – that is, security officers or IT managers of governmental agencies of Iran, although most of them had some expertise in both areas. At least %75 of them had eight years of experience in their fields. In the first round, 38 qualified people in the fields of security and e-government were chosen and the questionnaire was sent to them. 16 experts (more than one third) returned their completed questionnaires. In the second round, the questionnaire was sent to 16 experts who took part in the first round. Finally, 12 filled out questionnaires were received. The data were analyzed using SPSS. By following the various stages of e-government maturity, this study highlights the required security practices in each stage of the e-government maturity.

The research is based on Delphi technique that defined by Kaynak and Macauley as ''a unique method of eliciting and refining group judgment based on the rationale that a group of experts is better than (one) expert when exact knowledge is not available'' [30]. The Delphi method was conceived as a group technique whose aim was to obtain the most reliable consensus of opinion of a group of experts by means of a series of intensive questionnaires with controlled opinion feedback [31]. There are two logics behind this selection. First, this paper covers two distinct areas that include E.G and information security issues. There aren't enough experts who have good knowledge and experience in both areas. The power of this technique is based on the knowledge of the experts rather than the number of them. On the other hand, because of the iterative steps in this technique, the results are valid and acceptable. This method seeks to reduce the influence of factors such as group conformity, prestige, power and politics on the individual responses [32]. Finally, the experts reach consensus about the final results. There are indications, in the literature, that the minimum size of a panel of experts to involve in a Delphi exercise should be no less than 8 to 10 members [33]. In Delphi implementation, we restricted the number of rounds to two, in order to limit as much as possible the bandwagon effect and the weariness of respondents. After each round, the results along with their statistical analyses were returned back to the group members who sent new responses to the questionnaire; and the information was provided on the complete pattern of judgments in the first round so as to reduce this conformity pressure. Median ratings and IQR were calculated which depicted the experts' consensus to identify the major practices. In the next part of this paper, findings are presented.

4 Findings

According to the experts' opinions, the most important practices have been identified in each stage of e-government maturity and presented them in order of the stages. Finally, a comprehensive analysis was presented. As an indicator, the median within the range of ''Agree'' to ''strongly agree'' indicated that over 50% of participants agreed to choose it [28]. The factors, upon which no consensus was achieved, are highlighted in the tables.

4.1 First Stage: Cataloguing

The most important practices of information security management in the stage 1 are Physical information security control and Security culture establishment. As mentioned before, this stage is characterized by using traditional technology, and so physical information security control has significant importance as the results show. On the other hand, if someone wants to have a successful development of new phenomena, he/she should have comprehensive analyses of the current position. It can be observed that the factors related to these issues have the superior importance such as asset management, selecting and implementing information security management standard, determining information security policies and strategies, and human resource training. The experts didn't reach consensus about three practices: determining information security policies and strategies, incident management, and secure information system development.

4.2 Second Stage: Transaction

Based on the Median of experts' responses about the main practices of information security management in the stage 2 of the e-government maturity, the most important practices in this stage are database security, and legal aspects of organization e-presence. Because of bidirectional transaction in this stage, dynamic databases are needed, so that their security importance is increased. A unique standard, procedure and format are to be determined for communicating through networks for public usage. Accordingly, the manner of transaction will be changed. The organization is responsible for the received information from citizens and so obeys the legal consequences of this transaction. Other important practices are knowledge management of information security, physical information security control, security culture establishment, and audit management. The experts didn't reach consensus about three practices: asset management, incident management, and physical information security control.

4.3 Third Stage: Vertical Integration

In the vertical integration, a connection between local and state systems is created. For effective communication in this integrated environment, the similar information security policies and information security management standards should be determined. The employees of local and state offices should operate their tasks in a coherent manner; therefore designing a human resource program and setting a unified culture between them and also their training is essential. As the information flow within the organizations increased, we require more secure networks. The experts didn't reach consensus about one practice: physical information security control.

4.4 Fourth Stage: Horizontal Integration

Because of integrated functions through databases and information sharing, organizations become a one-stop shopping unit. As a result, everybody can access to any kind of information if there is no limitation. Hence, the importance of determining access level of information, information security organization, and database security increased. In information systems development, it is notable that there is no breach.

Furthermore, a problem occurring in one department causes another problem in another department, so not only having incident plans and good backups is necessary, but also existence of an accepted culture between all of the employees about information security is unavoidable. The other important practices are: information encryption, physical information security control and network security. The experts reach consensus about all of the practices.

4.5 Fifth Stage: Extension

Based on the literature, major attribute of this stage is personalized Web and so the Web interfaces are based on end users needs. From security perspective, everyone should do his/her specific tasks without interfering others. The information system attributes is changed into personalized security manner, but because the nature of information resources is similar to the previous stage, asset management is not so important. In contrast, considering the personalized security features in the process of information systems development is essential. Also it's evident that information encryption, security knowledge management and access level are essential, as well. Since human resources should perform different types of tasks, they need different security skills and the organization should have a comprehensive program for their training. For this diversity, having a common culture is not so important. The experts reached consensus about all practices.

4.6 Sixth Stage: Maturity

The most important stage in e-government maturity model from security perspective is the maturity stage. In this stage, the Internet and intranet have merged and the homepage is feeding online information from other institutions to the users. Further, the Web site is organized to solve problems and response to requests rather than presenting formal organizational structures and general information. Self-service is a key priority in this phase. By reason of increased importance of the Internet, the practices which are related to the network like network security, information transportation standard and information system security are very important. In this situation, everyone is regarded as a potential user of the system, until we can recognize his/her identity. So determining access level of information is imperative. Also, knowledge management and information security organization are important in this stage. Loading information in such a public place like Internet highlighted the degree of importance of incident management. Human resource training and security culture establishment received less importance comparing with the previous stage on account of self services for public users. The experts didn’t reach consensus about one practice: physical information security control.

4.7 Seventh Stage: Revolution

The revolutionary phase is characterized by data mobility across organizations, application mobility across vendors, and ownership to data transferred to the citizens. This is possible through intra and extra organizational mobility of data and services. The desired objective is to transfer data ownership and the orientation of database infrastructure to the end-users. The results show a significant decrease in the necessity of

security practices. It was happened by reason of the infrastructures and applications assigned to the users, and the organization is only a coordinator. Although there isn't consensus about physical information security control in some stages, the experts displayed strong agreement on its low importance in this stage. It might be refers to the lack of real physical organization and all of the interactions which occur in the virtual space. The practices related to the network and identification of people is important as the previous stage. The experts didn't reach consensus about one practice: incident management.

Moreover, to give a better vision which managers could pay attention to these practices in the future, we show the results in a new order. In the following charts, we categorize these practices into four charts based on their similar trends during the seven stages. In chart 1, there are six practices that have an increasing importance during the stages, but in stage 7 they loose their importance rather than the previous stage. As we discussed earlier, the special circumstances of stage 7 like assigning the applications to the users and the lack of real physical organizations cause to disability of organizations to control some aspects of information security. In addition to this lack of authority, the organization should be mature in some practices in 6 previous stages. In the stage 7, it is not necessary to focus on the practices that showed in the chart 1. In chart 2, there are five practices that reach their peak in stage 3 or 4, and then they loose their importance. Before stage4, each organization should reach internal maturity of e-government, so they should be mature in the practices that are related to the internal information security management. All of the practices that are shown in the chart 2 are inter-organizational factors. After stage 4, the organization practices focus on improving its relation with external organizations and customers, so these practices loose a part of their importance. In chart 3, there are 4 practices that their importance is ascending. As organizations begin developing their practices on the networks and extended use of public access networks, the items related to the network security and the tools that ensure the right people will have access to the right information becomes more important. In the chart 4, there are some practices that don't obey any special trend. They have an increasing importance during the initial stages, but they haven't any order in the second half of the stages.

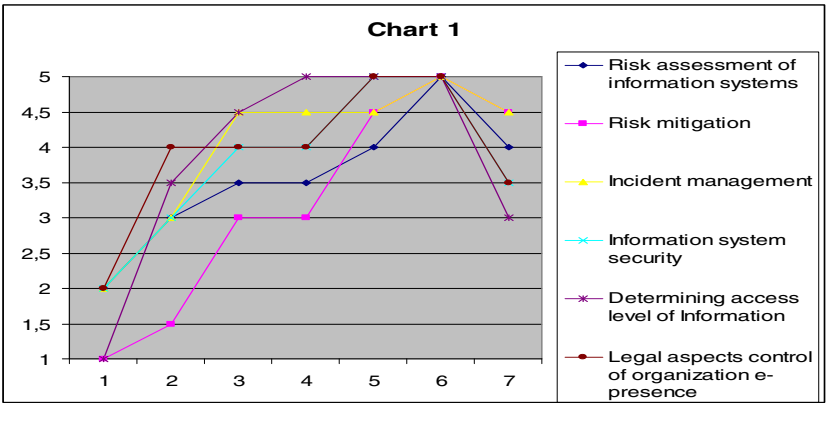


Chart 1. Six practices that have an increasing importance during the stages

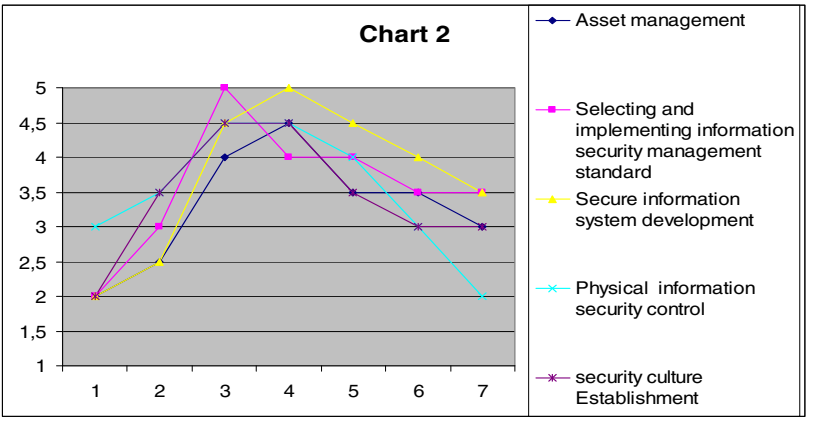


Chart 2. Six practices that reach their peak in stage 3 or 4, and then loose their importance

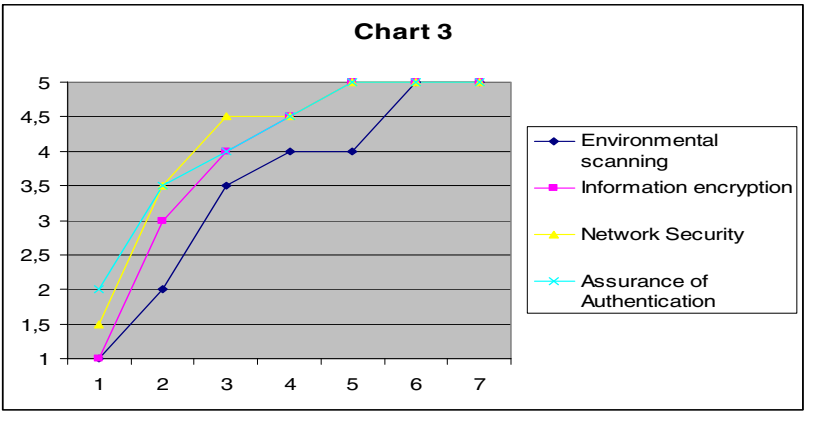


Chart 3. Three practices that their importance are ascending

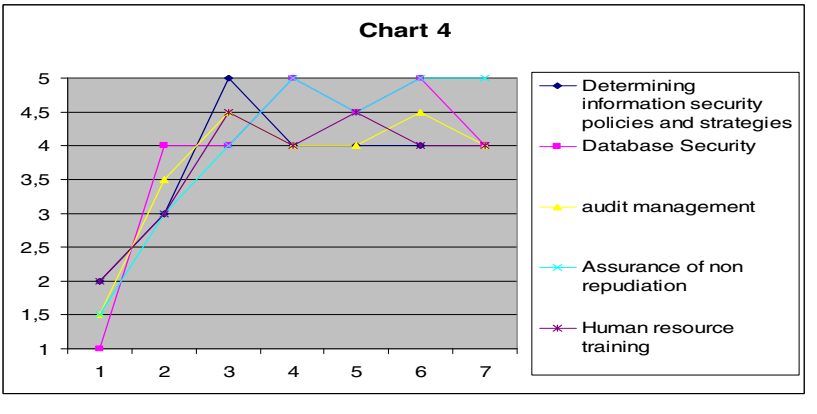


Chart 4. Some practices that don't obey any special trend

5 Conclusion

This paper tried to identify the importance of information security requirements during each stage of the e-government maturity. In the findings, the importance of twenty selected items was illustrated, and the most and least important practices in each stage based on the expert's opinions and the statistical data analysis were explained. The results can be used in the security policy making and security practices of governmental organizations according to their position in e-government maturity model and their special circumstances. As Hong suggested, we agreed that dealing with some limited aspects for information security management is not adequate and comprehensive, and we should draw on an integrated system of information security management [34]. In addition to this integration, the managers should have a precise insight about the requirements for traversing the path of their development. The findings of this study imply that information security is an ingredient of e-government that should fulfill the fundamental security properties of: availability, confidentiality, integrity, accountability, and information assurance. Therefore, proper information security policies, procedures and practices are vital for existence and evolution of e-government. On the other hand, each stage of e-government maturity needs special information security policies, procedures and practices. Future researchers can develop the twenty practices by considering their special situation and combine the results of this paper to the missions of their organizations. By emerging new modern technology and promoting e-government services, it will be useful that the special information security requirements be extracted and identified.

References

1. Mitrou, L., Karyda, M.: Employees_ privacy vs. employers_ security: Can they be balanced? Telematics and Informatics 23, 164–178 (2006)
2. Davies, P.: Constructing electronic government: the case of the UK Inland Revenue. International Journal of Information Management 25, 3–20 (2005)
3. Fu, J., Farn, C., Chao, W.: Acceptance of electronic tax filing: A study of taxpayer intention. Information & Management 43, 109–126 (2006)
4. Sarathy, R., Muralidhar, K.: Secure and useful data sharing. Decision Support Systems 42, 204–220 (2006)
5. Jaeger, P.T.: Deliberative democracy and the conceptual foundations of electronic government. Government Information Quarterly 22, 702–719 (2005)
6. Chou, T., Chen, J., Pu, C.: Exploring the Collective Actions of Public Servants in e-government Development. Decision Support Systems 45(2), 251–265 (2008)
7. Parent, M., Vandebeek, C.A., Gemino, A.C.: Building citizen trust through e-government. Government Information Quarterly 22, 720–736 (2005)
8. Torres, L., Pina, V., Acerete, B.: E-government developments on delivering public services among EU cities. Government Information Quarterly 22, 217–238 (2005)
9. Layne, K., Lee, J.: Developing fully functional E-government: A four stage model. Government Information Quarterly 18, 122–136 (2001)
10. Anderson, K.: Convergence: A holistic approach to risk management. Network Security 5, 4–7 (2007)
11. Whitson, T., Davis, L.: Best Practices in Electronic Government: Comprehensive Electronic Information Dissemination for Science and Technology. Information Quarterly 18(2), 79–91 (2001)

12. Gunter, B.: Advances in e-democracy: overview Aslib. Proceedings: New Information Perspectives 58, 361–370 (2006)
13. Tsai, N., Choi, B., Perry, M.: Improving the process of E-Government initiative: An in-depth case study of web-based GIS implementation. Government Information Quarterly 26, 368–376 (2009)
14. Ebbers, W.E., Dijk, V.: Resistance and support to electronic government, building a model of innovation. Government Information Quarterly 24, 554–575 (2007)
15. General Accounting Office, Electronic Government: Challenges Must Be Addressed with Effective Leadership and Management. General Accounting Office, Washington, D.C. GAO-01-959T, July 11 (2001)
16. Anderson, B.: Security and the future: Anticipating the event of terror. Geoforum 41, 227–235 (2010)
17. Von Solms, B.: Information Security – The Fourth Wave. Computers & security 25, 165–168 (2006)
18. Wiant, T.L.: Information security policy's impact on reporting security incidents. Computers & Security 24, 448–459 (2005)
19. Von Solms, R.: Information security management: why standards are important. Information Management & Computer Security 7, 50–57 (1999)
20. Gupta, A., Hammond, R.: Information systems security issues and decisions for small businesses: An empirical examination. Information Management & Computer Security 13, 297–310 (2005)
21. Dias, G., Alo Paiva, R., José, A.: A simple model and a distributed architecture for realizing one-stop e-government. Electronic Commerce Research and Applications 6, 81–90 (2006)
22. Mitchell, R.C., Marcella, R., Baxter, G.: Corporate information security management. New Library World 100, 213–227 (1999)
23. Hutter, D., Mantel, H., Schaefer, I., Schairer, A.: Security of multi-agent systems: A case study on comparison shopping. Journal of Applied Logic 5, 303–332 (2006)
24. Halchin, L.E.: Electronic government: Government capability and terrorist resource. Government Information Quarterly 21, 406–419 (2004)
25. Gil-Garcia, J.R.: Information technology policies and standards: A comparative review of the states. Journal of Government Information 30, 548–560 (2004)
26. Lambrinoudakisa, C., Gritzalisa, S., Dridib, F., Pernu, G.: Security requirements for e-government services: a methodological approach for developing a common PKI-based security policy. Computer Communications 26, 1873–1883 (2003)
27. Elahi, S., Abdi, B., Shayan, A.: E-banking and managerial challenges: change management. In: First annual e-banking summit, Tehran, Iran (2007)
28. Tsaura, S., Linb, Y., Linc, J.: Evaluating ecotourism sustainability from the integrated perspective of resource, community and tourism. Tourism Management 27, 640–653 (2006)
29. Sackman, H.: Delphi Critique. Rand Corporation (1975)
30. Kaynak, E., Macauley, J.A.: The Delphi technique in the measurement of tourism market potential: The case of Nova Scotia. Tourism Management 5, 87–101 (1984)
31. Landeta, J.: Current validity of the Delphi method in social sciences. Technological Forecasting & Social Change 73, 467–482 (2006)
32. Deshpande, M., Aniruddha, S., Richard, N., Prakash, M.: Metadata-driven Delphi rating on the Internet. Computer Methods and Programs in Biomedicine 77 (2005)
33. Mitchell, V.M.: The Delphi technique: an exposition and application. Technol. Anal. Strateg. Manag. 3, 333–352 (1991)
34. Hong, K.S., Chi, Y.-P., Chao, L.R., Tang, J.-H.: An integrated system theory of information security management. Information management and computer security 11, 243–248 (2003)

Dynamic Device Configuration in Ubiquitous Environments

Abdullahi Arabo, Qi Shi, and Madjid Merabti

Distributed Multimedia Systems and Security (DMSS) Group
School of Computing & Mathematical Sciences
Liverpool John Moores University, Byrom Street, L3 3AF, UK
{a.arabo,M.Merabti,Q.Shi}@ljmu.ac.uk

Abstract. The need of devices in crisis management to be configured dynamically by detecting device characteristics i.e. polices defined by the user or organization, contextual information relevant to a given scenario etc is of paramount importance. Where such information can either be already predefined or the user is allowed to define such information via automatically generated input User Interface (UI) associated to one or more extensible markup languages. Hence, the layout of the devices and behavior will be automatically configured based on policy settings and contextual information in a dynamic manner as such information changes. We present a method that allows dynamic configuration of devices that improves information systems flexibility via realizing dynamic configuration of components and enhancing management and functionality of such devices and security issues within the environment. Moreover, as this method will provide an instant configuration of devices at runtime, the components can provide and be used with uninterrupted running ability.

Keyword: Dynamic Policy specification, Identity Management, Crisis Management.

1 Introduction

Due to technological development we now attained the point whereby electronic devices are customary in every aspect of our life. Today, we encounter numerous mobile devices within emergency crisis management environments' as well as within the public domain i.e. home and office environment. The notion of Ubiquitous Computing coined by *Weiser* [1] has received increasing attention as a result increase adaptability of such technology. Mobile Ad-hoc Networks (MANets) form one of the fundamental building blocks for ubiquitous computing environments and are increasingly used to support mobile and dynamic operations such as emergency services, disaster relief and military networks. However, this growing trend demonstrates the significance of dynamic device configuration at runtime with limited or uninterrupted ability of service provision within crisis management scenarios. As well as providing information system flexibility, where the quality of adapting to changes which reflects the adaptability to requirements, context and environments, these can be in terms of market-oriented, development-oriented, application-oriented and technology-oriented flexibility[2, 3].

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 263–273, 2010.

Hence, devices in crisis management can be configured automatically by detecting devices characteristics, polices, contextual information relevant to a given scenario etc. Where such information can either be already predefined or the user (target device) is allowed to define such information via automatically generated input UI associated to one or more extensible markup languages. Hence the layout of the devices and behavior will be automatically configured based on policy settings (polices can be as simple or as complex as the need of the user) and contextual information.

This will also ascertain the overall functionality of the devices within the network throughout its connection to the server i.e. if one of the policy elements prohibits the user of the devices from performing certain action like receiving or sending highly confidential information, the user of the devices will not be presented with such functionality. Another example will be where certain functionality of the devices is based on the availability of some contextual information. Without such information the functionality will not be available. The devices configuration is expected to response to rapid change in contextual information dynamically to meet the need of the scenario and system.

Considering the nature of most emergency or crisis management scenarios and the rapid proliferation of interaction between systems from various teams or disaster recovery and planning organizations, dynamic device configuration is of paramount importance. Hence, in this paper we focus on dynamic configuration based on devices distinctiveness, user profiles, polices and contextual information in large scale and dynamic crises of natural and man made disasters. Natural disasters such as fire, floods, volcanic eruptions, earthquakes and tornados are mostly on large scales and dynamic, and require the involvement of various emergency services and organisations.

The rest of the paper is structured as follows: related work is discussed in section 2, in section 3 we described the process involved in dynamic device configuration; where the section is further divided in three sub-sections UI configuration, contextual information configuration and data sharing. Section 4 presents a scenario that our proposed methods and anticipated implementation is based upon, provides further information in regards to our current stage of implementation. Section 5 provides some evolution of our proposed method and algorithms. We conclude the paper in section 6, where we also pointed out areas of future research work within the area.

2 Related Work

Current device used within the community are complex for a lay user, hence it is very difficult for users to configure and use these devices in disaster emergency situation, where the use of the devices is crucial and need to be configured dynamically as the situation changes. Current research in this area and Service-Oriented Architecture has failed to produce convincing results for seamless integration between devices. Frameworks such as OSGi [4, 5], UPnP [6], DPWS ,DLNA [7], HAVi [8] and ePerSpace [9] are used for integrating home Networked Appliances [10]. However, these solutions required the user to configure these devices manually and in some solutions are managed via centralized providers or fixed configuration files. Services are usually discovered and composed using middleware protocols and interoperability issues are addressed using agreed standards. However, the current standard is not capable of addressing all these issues. This is an extension to the current traditional devices

configuration based on the UPnP [6], where an existing configuration file is used to configure the devices, the new addition is the (UI) which allows users to input the relevant policies of the user depending on its circumstances and environments', also part of the ability of the user either using pres-defined profile information or supplying such information via a dynamically generated UI.

The solutions mentioned above do not provide any mechanism to automatically and dynamically discover and compose devices and services. Compositions are based on application based serialization. Such services become more heterogeneous and managing such a framework will be more complex where the amount of control placed on device and service integration becomes more difficult. Different service providers use different communication standards and middleware, due to which interoperability becomes a main problem requiring more sophisticated solutions. Hence, there is a need for a new architecture and methods to filter through the restrictive proprietary of existing solutions. This problem has been addressed, in part, using peer-to-peer (P2P) [11] technologies whereby digital content can be distributed and discovered using global communications [12]. In this paper, we proposed new methods on how to dynamically configure devices in emergency situations automatically without requiring much technical knowledge of the user of the device.

3 Device Configuration

Before going into the details of device configuration, let's define some of the terms used within this paper:

- Target or target device: this connote to the device or resource that other devices/resources will like to communicate with i.e. by either requesting profile information, data etc.
- Subjects: this denotes to either a devices or service that is requesting information from the target device.
- Proxy: this signifies to subject devices or services that the subjects can go through to request information from the target devices. Or the target device will use to send relevant information to the subject(s). For example *target A* is connected to *subject B*, *B* is connected to *subject C*, but *A* has no direct connection to *C,* therefore *A* is indirectly connected to *C* via *B,* hence *B* is the proxy node/subject.

It also worth pointing out that although we have utilize the components within the proposed methods as either devices or persons (individuals) representing different organization holding devices, it can also be equipments i.e. medical equipment, fire service equipments, other equipments/service that might help users of the devices to fulfill their responsibilities in large dynamic crisis management ubiquitous environments.

We divide the process of dynamic device configuration in three main stages (refer to listening 1) as follows: configure the interface based on either pre-defined or user defined policy, obtain contextual information within the range of the device and configure it dynamically based on the contextual information gathered and negotiate policies and user request dynamically.

A. UI configuration

Each device will have some default settings of device profile, user profile and policy settings. When the device is introduced into the environments the automatic configuration process is initiated. Fig 1 presents the flowchart of the required process. If user preferred settings are automatic detected a user will have the option to either modify such settings or not. The processes also neither applies if there are no predefined user settings. If default settings are selected, the information is then presented to the module that automatically and dynamically configures the device.

Step-by-step Device configuration (Listening I)

```
Configure device interface
List of nodes within range
Compare target and subjects polices
        if match
            Create list of nodes
            Present list to device
            Configure device
          Repeat is context changed
        End if
            If any proxy node
                Inform target device
                Create list of proxy subjects
            End if proxy
        Else
          Provide alternative policy to be    negotiated
        End if match
```

Then the configured devices should be presented to the user. If the user opted to modify the default settings or their own specified settings then an interface will be provided for them to modify such information. After such modifications the process of configuration will start.

Initially the required configuration settings are detected (see Fig 2) from the devices involved, if such information defined by the user exist or the default settings are found the user is giving the option of either modifying the settings or not. The next process involves which settings need changing and this information is use in the next step to configure the device in terms of policies requirements, user profiles etc. After a successful configuration, the configured device is then presented to the user. This process is done dynamically and automatically, if the contextual information of the devices has change only such information can be changed dynamic with little or no impact to the availability of service and functionality of the device to the user.

This can also be implemented in a way that the user is permitted to specify one or more preferences associated with the device configuration settings, including policy, some parts of user profile etc. As well as defining new fields to be used for data entry in terms of new properties, for policy creation or user profile specification. The layout of such input interface will be based on a Meta data i.e. using XML files to generate data entry forms dynamically. Some of the computing devices we anticipate to be used are designed to interact in a client server setting. Expected Client devices includes PDA's, mobile phones, palmtop etc. The communication between such devices and the server is done wirelessly through communication interfaces, communication can be via audio, text, pictures, images and data transfers such as files between devices or client and server. Each mode of communication will be automatically detected and determined based of devices setup using policies and relevant contextual information provided and the given scenario.

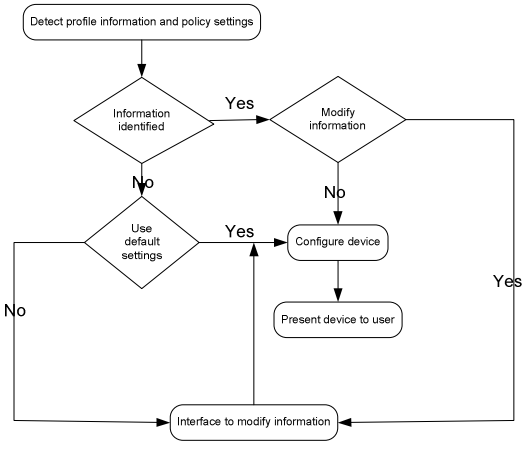


Fig. 1. Flow chart for configuration device

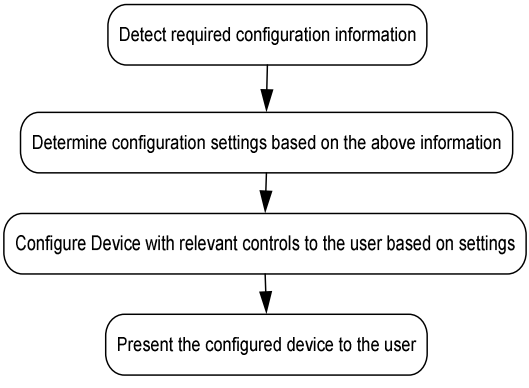


Fig. 2. Steps for dynamic configuration

B. Contextual information configuration

When a device is first registered within a network it will be able to obtain the list of devices within its range, listening 2 specifies the information that will be passed on the connecting device. As shown in Fig 3, the devices within the circle are the once that meets the requited policy of the device and that are within the range. Three other subjects are shown within the figure in a red color. Two of the subjects are within the range but failed to meet one or more of the policy requirements while one of the subjects has not fulfilled any of the policy requirements. The two subjects within the range of the device might have also change their profile type, which leads to a change of context, this information is relayed to the target devices dynamically. By proxy nodes we are referring to other nodes that are inter-connect with the nodes that are within the target device but are not within the target device range itself- hence not visible to target devices but visible to other nodes within the target device rang – as

depicted in Fig 4. These proxy nodes can either fulfill target device policy requirement or not. Some of the issues of concern are what happened to information that you have send to a node within your range, should the node be able to pass on such information to your proxy node. This issue in particular will be covered in the next stage of device configuration.

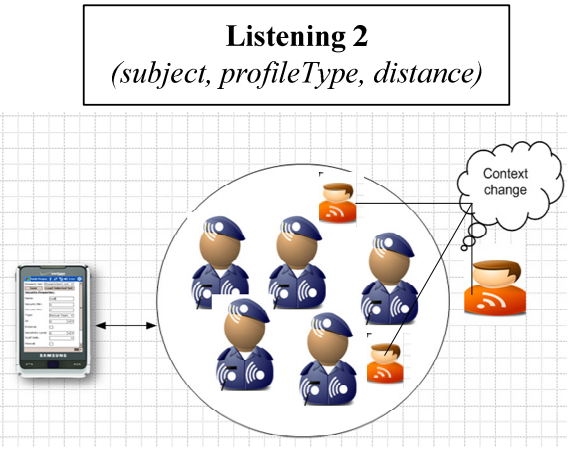


Fig. 3. Dynamic Context Configuration

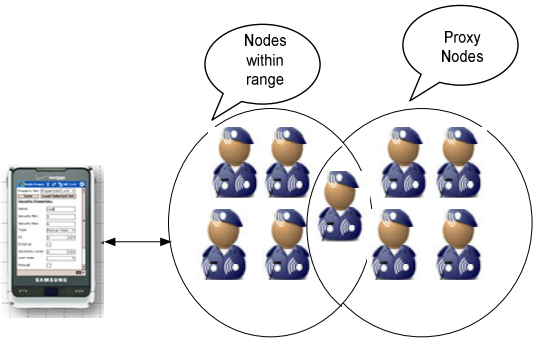


Fig. 4. Proxy Nodes

C. Data sharing

When a subject (that fulfils the requirement of the target device as explained in the sections above been either a direct node or proxy node) needs to request certain information from the target devices. The request need to be done either directly or via proxy node based on the energy matrix [13]. The target device will response with the information provided in listening 3.

The condition specifies what the subject can do with the information received i.e. can they pass on to the other devices, can they modify it, a condition of automatically deleting the information from the devices after certain policy elements of the subject or the target has been changed i.e. leaves the environment etc. In summary the condition is a policy file of rules/ stipulations of information usage. Hence, this gives the target device full control of their information and eliminates the problem of building full target profile from sub profiles either by subjects or Sybil attacks [14] or man-in-the-middle attacks [15] that can lead to the leakage of information during or after communication.

Listening 3
Target, Subject, Condition

D. Dynamic Policy Specification

To handle such situations as well as enable user to be in total control of not only their partial identities but any information they may have shared with other users. We proposed a new dynamic policy specification that is simple and gives users more control.

By policy we mean a set of rules that can be used to determine if a given query will materialize or not. The policy consists of a head (the left hand part of the policy) and body (the right hand part of the policy). Initially to deduce the head of the rule, all the body predicates must be deducible in such a way that the constraint is also satisfied. Queries can be in two forms: perform an action or request a credentials or information.

Our policy specification make use of dynamic literal so as to meet the demanding nature of ad-hoc environments rather than just hard coding the rules in the rule engine.

$$Dynamic\ Literals\ \langle L \rangle ::= A \pm \{P(\overline{X})\}$$

$$Dynamic\ Rules\ \langle R \rangle\ ::= A \leftarrow \overline{L}$$

Preliminaries: the notation of a line above a variable represents a (possible empty) sequence of distinct variables. $P \in Pred,\ A \in Atom$. This definition is based on first order function-free $\Sigma = (Const, Pred)$ with possibly infinite many constants $Const$ and a set of finite predicates names $Pred$. This generates a set of atoms which represents predicates names applied to expirations of appropriate rules, denoted as $Atom$. A policy P is a finite set of rules R . We use predicates for variety of purposes, dynamic predicates names represents actions (or access request), and dynamic rules define the conditions and state updates associated with actions, hence dynamic rules are also called *action definition*.

$$DelInfo(u, f) \leftarrow Rang(u_1, u_2)$$

The policy above represents a simple example of deleting information based on the function for the range between two devices, the owner of the information u_1 and the user who already have the information u_2. This policy can be further expanded to check if the user U has the file with $has(u, f)$.

$$\therefore DelInfo(u, f) \leftarrow Rang(u_1, u_2), has(u, f)$$

A user of profile information or any information can also specify who is allowed access, read, modify etc the information.

$$canRead(u, f) \leftarrow \neg Sealed(u, f)$$

The policy above states that information f can be read or accessed by user u only iff the information is not sealed as un-readable or accessible for the user u.

4 Implementation

The implementation of the above proposed methods is still in its early stage. We have only implement part of the first stage of the dynamic configuration presented above and have testing some of the methods in terms of access control when sharing data between devices of various organizations either from the same or different organization. Full details of the access control implementation of data access control can be found in [16].

However, in order to understand problems that requires dynamic device configuration, assigning users control of their information as well as been able to define policies, request or share information with proxy or directly connected subjects based on contextual information and potential access issues, a rudimentary model/scenario is presented, readers are referred [16] for details of the scenario presented in this paper .

Consider a situation involving a set A of n organizations, where $A = \{a_1, a_2 \ldots a_n\}$. These organizations might be emergency services, hospitals, companies, volunteer groups and so on. These organizations can also be or may be split into sub-units. Thus the police may for example be represented by more than one element from A.

We assume that all members of an organization have access to data within that organization. Conversely, as we have seen earlier on, in a crisis situation this may not be the case. Consider some organization a_i and piece of data d. The organization a_i forms part of a network made up of elements of A. Let $A' \subseteq A$ be all organizations downstream from a_i; that is, all organizations reachable in the network from a_i.

As presented in Fig 3, all the subject shown in the diagram can be from different organization that have fulfill the required policy of the target device say a police device. Hence, for any of this subjects to further access any other information from the target devices the request should be done either directly or via proxy subjects. In Fig 4, assuming the devices are representatives from various organization. A subject with the nodes within range (say an ambulance crew) want to access information about say a police offer to take report from the victim been treated, but the office only appears as part of the subjects in the proxy node devices range. These communication as it is of crucial importance need to be done via the proxy subject that has a direct connection with the required subjects that hold the information d that is required by the original target devices (ambulance crew). The complication arises in a situation where by the intermediary device is not allowed to access the data d send. Hence, the condition send attached to the information request as depicted in Listening 3, which will be able to prevent the subject in the middle of this communication from viewing, editing and passing on this information to any other devices but only to the requesting target device.

5 Evaluation

The method presented in this paper provides a number of advantages against existing traditional methods. It allows devices to be configured during runtime with little or no interruption to service provision. Users are able to make some changes to the settings, which will instantly take effect dynamically, it makes the configuration process dynamic and increases the range and security control of the application and eliminates the needs of an expert to change policies and other settings by allowing users to define such policies via a simple UI (to be implemented), which will in turn be translated into policies and the device get configured dynamically accordingly. As it currently stand we only anticipate problems when it comes to dealing with more complex policy negotiations, which might required a more in depth method and algorithm to handle such situations and this forms part of our future direction of study as well as enabling multi hop communication and other security issues such as integration of services. Where to access one service you need to go through other services which might have not meet the target device policy requirements.

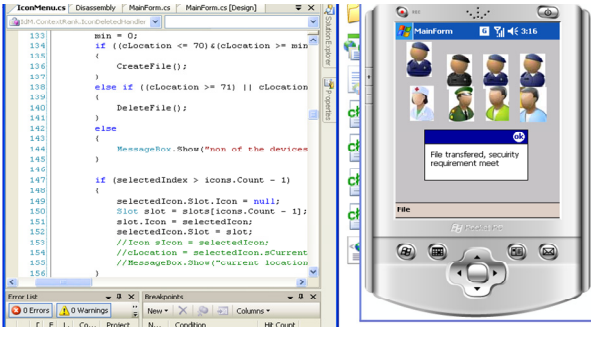


Fig. 5. Transfer File- Policy Fulfilled

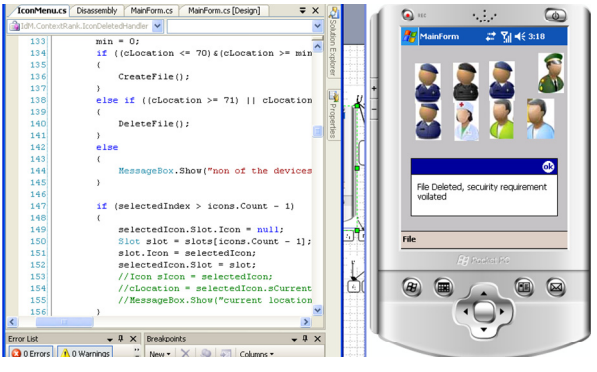


Fig. 6. Delete File-Policy Violated

The implementation of the above example is shown in Fig 5 and 6, where profile files are either transferred or deleted based on the location of the owner of the information and its location in relation to other users, contextual information and policy specification. Where each device is configured dynamically to according to the policy specified and re-configuration is also done dynamically and automatically after any change of contextual information, policy etc.

6 Conclusion

We have outlined a practical implementation of dynamic device configuration in ubiquitous environments at runtime with an uninterrupted functionality of the systems or devices. This process is divided into three main stages. Details of each stage have been analyzed and relevant methods are presented where appropriate. Details of our current implementation stage has been present and we have evaluated our proposed methods based on our early implementation results with regards to other related similar work and proposed methods. Future work involves the full implementation of proposed method and solution.

References

1. Weiser, M.: The computer for the 21st century. ACM SIGMOBILE Mobile Computing and Communications Review 3(3), 3–11 (1999)
2. Judith, G., Franz, S.: Information System Flexibility and the Performance of Business Processes. Journal of the Association for Information Systems (JAIS), 1–43 (2005)
3. Lu, L., Zongyong, L., Ruibo, L.: Improving Information System Flexibility through Remote Dynamic Component Configuration. In: 2006 International Conference on Service Systems and Service Management 2006 (2006)
4. Lee, J.-J., Huang, C.-Y., Lee, L.-Y., Lei, C.-L.: Design and implementation of secure communication channels over UPnP networks. In: 2007 International Conference on Multimedia and Ubiquitous Engineering (MUE 2007), pp. 307–312 (2007)
5. OSGi: The OSGi Service Platform - Dynamic services for networked devices (2008), http://www.osgi.org (cited 2005)
6. Lee, J.-J., et al.: Design and implementation of secure communication channels over UPnP networks, Institute of Electrical and Electronics Engineers Computer Society, Piscataway, NJ 08855-1331, United States, Seoul, South Korea (2007)
7. DLNA: DLNA Overview and Vision Whitepaper (2006), http://www.dlna.org/en/industry/about/dlna_white_paper_2006.pdf (cited 2008)
8. HAVi: The HAVi Specification (2004), http://www.havi.org/HAVi_1.1.pdf (cited 2006)
9. ePerSpace: Towards the era of personal services at home and everywhere (2005), http://www.ist-eperspace.org/ (cited 2005)
10. Merabti, M., et al.: Managing Distributed Networked Appliances in Home Networks. Proceedings of the IEEE 96(1), 166–185 (2008)
11. Mol, J.J.D., et al.: Free-riding, fairness, and firewalls in P2P file-sharing. IEEE, Piscataway (2008)

12. Li, J.: On peer-to-peer (P2P) content delivery. Peer-to-Peer Networking and Applications 1(1), 45–63 (2008)
13. Abdullahi, A., Qi, S., Madjid, M.: Situation Awareness in Systems of Systems Ad-hoc Environments. In: Proceedings 5th International Conference Global Security, Safety, and Sustainability, ICGS3 2009, London, UK, September 1-2, vol. 45, pp. 27–34. Springer, Heidelberg (2009)
14. Haifeng, Y., Michael, K., Gibbons, P.B., Flaxman, A.D.: SybilGuard: Defending Against Sybil Attacks via Social Networks. IEEE/ACM Transactions on Networking 16(3), 576–589 (2008)
15. Benjamin, A., Geoff, H.: Detecting Man-in-the-Middle Attacks by Precise Timing. In: Third International Conference on Emerging Security Information, Systems and Technologies, SECURWARE 2009, pp. 81–86 (2009)
16. Bo, Z., Abdullahi, A., Oliver, D., David, L.-J., Madjid, M., Qi, S., Adrian, W., Rachel, C., Glyn, J., Arnold, K.L.Y.: Data Flow Security Analysis for System-of-Systems in a Public Security Incident. In: The 3rd Conference on Advances in Computer Security and Forensics (ACSF 2008), Liverpool, UK (2008)

Mitigation of Control Channel Jamming via Combinatorial Key Distribution

Abolfazl Falahati and Mahdi Azarafrooz

Iran University of Science and Technology, School of Electrical Engineering
Tehran, Iran
afalahati@iust.ac.ir,azarafrooz@ieee.org

Abstract. The problem of countering control channel jamming against internal adversaries in wireless ad hoc networks is addressed. Using combinatorial key distribution, a new method to secure the control channel access is introduced. This method, utilizes the established keys in the key establishment phase to hide the location of control channels without the need for a secure BS. This is in obtained by combination of a collision free one-way function and a combinatorial key establishment method. The proposed scheme can be considered as a special case of the ALOHA random access schemes which uses the common established keys as its seeds to generate the pattern of transmission.

Keywords: Combinatorial designs, Wireless multiple access, Control channel accesss security.

1 Introduction

Multiple access based on orthogonal frequency division multiplexing, or (OFDMA) provides a unifying framework to allow users share the wireless media [1]. In order to fulfill the channel allocation and routing functions efficiently, a necessary overhead in the form of exchanging messages known as the control channel are needed. However from a security point of view, a pre-assigned control channel introduces a single point of failure. In fact, an adversary can perform a denial-of-service (DoS) attack [2,3] by jamming the pre-assigned control channels which requires much less energy than in fulltime jamming [4]. The traditional spread spectrum anti-jamming strategies are not effective in the case of internal attackers since they have access to DSSS or FHSS sequences. Generally not so much works have been done on counteracting the problem of link layer jamming.

Within our knowledge, three major works have been carried out to address the control channel jamming problem. This was first addressed by Chan et. al. [4] in the context of GSM networks. Their proposed approach made a use of keyed hash functions to locate the control channels. In fact they used coding theory to develop a key distribution scheme (BBK assignment) that guarantees the resilience and identification of internal adversaries by assuming that no more than a fixed maximum number of traitor users exist in the network. In the second work [5] Tague et. al. proposed a framework for control channel access

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 274–281, 2010.

schemes using probabilistic assignment of cryptographic keys to users so that nodes can discover the location of the control channels in time/frequency with certain probability. Their method allows for graceful degradation in the control channel secrecy as a function of the number of compromised nodes, as opposed to the threshold approach in [5].

In a different approach [6] Lazos et. al. proposed a scheme that unlike two previous works, doesn't need a secure BS and is implemented in an ad-hoc network. In fact, they proposed a randomized distributed scheme that allows nodes to establish a new control channel using frequency hopping. Their method differs from classic frequency hopping in that each node follows a unique hopping sequence in a way that no two nodes share the same hopping sequence which in turn mitigates the impact of node compromise.

In the our work we propose a novel scheme that enables the nodes to exchange their control channel messages securely with their one hop neighbors eliminating the need to a secure BS.

2 The Employed Model and Preliminaries

2.1 Combinatorial Key Establishment

A set system is a pair (X, A) where the elements of X are called points and A is a finite set of subsets of X called blocks. The degree of a point $x \in X$ is the number of blocks containing x. (X, A) is regular (of degree r) if all points have the same degree, r. If all blocks have the same size k, then (X, A) is said to be uniform (of rank k). A (v, b, r, k)- design is a set system (X, A) where $|X| = v$, $|A| = b$, that is uniform of rank k and regular of degree r.

In combinatorial key establishment scenarios [7, 8], each block of the system is associated with a node in the sensor network. In fact, the points in the block are the key identifiers of the keys given to the corresponding node. In our work, we use the Lee-Stinson approach. For a thorough studying of this method refer to [8].

2.2 Control Channel Access Model

We use the same OFDMA framework for multiple access protocol as in [5]. We let $C = \{C_1, .., C_j, C_M\}$ denote the set of orthogonal carriers. We assume that time is slotted denoted by t and that an initial portion of each time slot is specified by $S = \{S_0, ..., S_{s-1}\}$ dedicated to control messages.

3 Proposed Scheme

In the proposed scheme we use the combinatorial key pre-distribution scenario to address the secure channel assignment problem. We also use Noubir scheme [11] to provide a collision free one-way function for assigning channels reliably. First, we apply the Stinson-Lee approach to the network to provide every node

with some keys. Using these pre-distributed keys, every node can independently determine if they share a common key with its neighbours in $O(1)$ time. The sensor nodes then use these common keys to make identification codes to communicate with each other reliably and securely using the suggested scheme. In fact, using the combinatorial key establishment helps us to control the channel interferences while maintaining the secure characteristic of the scheme.

Assume that two neighbor nodes N_i and N_j are going to communicate in a secure channel reliably. If N_i and N_j share a common key then utilizing this common key a polynomial of degree d over a Galois field $GF(p^q)$ is constructed as the identifier of their transmission (Algorithm I). Assuming the number of nodes is N, this is possible if $(p^q)^(d+1) \geq N$. As before, let $C = \{C_1, .., C_j, ..., C_M\}$ be the set of orthogonal channels. Furthermore let $M = p^q$ be the number of orthogonal channels. The polynomial associated with the nodes is $I(c) = \sum_{k=0}^{d} I_k c^k$ where c is an element of C. This polynomial is then evaluated over $|S|$ elements of $GF(p^q)$ where $|S|$ is the number of sub-time slots. Since two polynomial of degree d cannot be equal over more than d points without being identical, then if $|S| > (N-1)d$, there exist at least one point where a given transmitter differs with all the other transmitters polynomials evaluation. If the number of sub-slot times $|S|$ divides $p^q - 1$, by shifting the transmission frames we still have a polynomial of degree d and packets are transmitted in slot k by using $I(w^k)$ orthogonal channel (where w is a root of unity of order $|S|$), then transmission of two arbitrary node can still not intersect over more than d slots without being equal. This is because, the evaluation of this polynomial is now a Reed-Solomon code-word which is cyclic. However by shifting the set of possible polynomials is reduced by a factor of $1/|S|$. This introduces the new constraint $(p^q)^{d+1} \geq |S|N$. Some special guard time equal to sub-slot duration is also required to avoid loosing the properties of the used polynomials. This will keep the polynomial property of bounded collisions between transmissions valid. Thus transmitters will collide at most over d slots. Finally, guard times have to be inserted between sub-slots to avoid packet collision for shifts less than a sub-slot duration. This results in the following condition $(p^q)^{d+1} \geq 4|S|N$.

The proposed scheme can be considered as a special case of the ALOHA random access schemes. $I(w^k)$ can be considered as random number generated using a the shared key between nodes. However, given the properties of this random number generation, we can guarantee that the probability of collision is zero after a pre-determined number of re-transmissions.

In terms of secrecy, listening to the complete pattern of transmissions of a nodes during a time slot even with knowledge of the keys doesn't leak useful information about the location of secure channel since neither adversary nor the node itself will be able to determine the location of collision free channels before transmissions. The only thing the transmitter node knows is that there is at least one collision free channel that a receiver node which has a common key within its range could receive the message by listening to it. Therefore the best strategy an adversary can pursuit is to jam a specific channel from the beginning of a time slot. Although our suggested scheme requires transmission in multiple channels

during a time slot and is not efficient in terms of communication complexity but as far as we know this is the only key establishment based channel allocation for ad hoc networks which can fulfill the security and reliability tasks of a secure control channel allocation scenario.

Algorithm 1. Selecting the secure and reliable channel between nodes N_i, N_j

Setup: Selecting the appropriate parameters $(r, M, |S|, d)$ using the Table I and Constraint Eq. 4
Use Lee-Stinson key distribution approach.
if $K_{N_i} \cap K_{N_j} = \emptyset$
  N_i and N_j cannot find a channel
else
  $I \leftarrow (K_{N_i} \cap K_{N_j})$
  $(I_0 I_1 ... I_d) \leftarrow I$
  $I(c) = \sum_{k=0}^{d} I_k c^k$
  $s \leftarrow 0$
  $w \leftarrow$ root of unity of order $|S|$
  while sub-timeslot $s < (2|S|)$ do
    Send and Receive Control information on $(frequency, time) = (I(w^{s/2}), s)$
      $s \leftarrow s + 2$
end

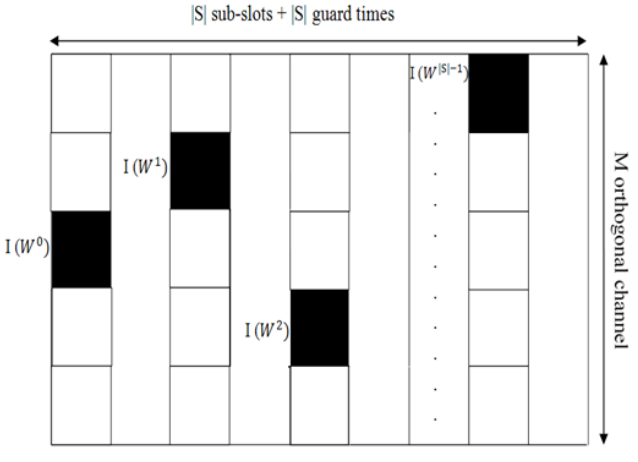


Fig. 1. Schematic pattern of collision free transmission in second scheme

3.1 Numerical Results of the Scheme

The main problem with our proposed method is that the large size of WSNs N makes the number of required time slots and orthogonal channels impossible to

be provided. Nonetheless the effective number of nodes that are causing interference will be less than N. Suppose that N_i and N_j are two nodes that are in each other's neighbourhood. The probability that N_i and N_j share a common key in Stinson-Lee approach is $p_1 = \frac{k}{r+1}$ where k is the number of keys per node and r is the number of nodes per key. Furthermore the probability that the common key between N_i and N_j is repeated in two other nodes is $p_2 = \frac{1}{kr}$ where kr is the size of key pool. We also suppose that the interference range to be less than $2R$ where R is the communication range. Therefore the $N_{effective}$ will be $p_1 p_2 N p_{intf}$ where p_{intf} can be derived using the fig 2 as follows:

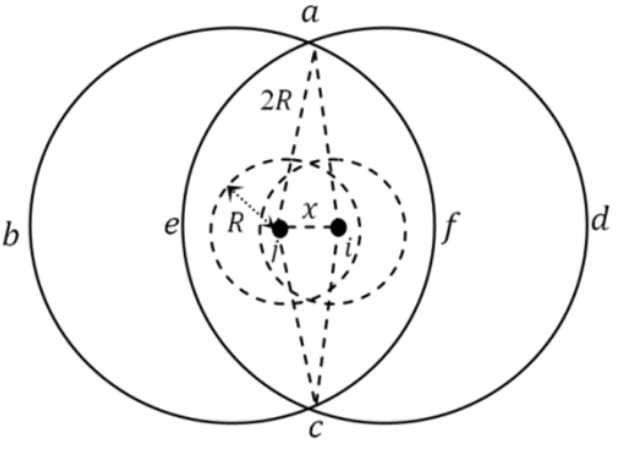


Fig. 2. The interference region

As illustrated in Fig 2, the area of the overlapped region $aecf$ is:

$$S_{aecf} = 8R^2 \cos^{-1}(\frac{x}{4R}) - x\sqrt{4R^2 - \frac{x^2}{4}} \tag{1}$$

Let the distance between i and j is $x(0 \leq x < R)$. The probability distribution function of x is given by $F(x) = P(distance < x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}$. Thus, the probability density function is $f(x) = F'(x) = \frac{2x}{R^2}$. The expected area of the overlapped region S_{aecf} is given by:

$$\bar{S}_{aecf} = \int_0^R S_{aecf}(x) f(x) x \approx 3.1565\pi R^2 \tag{2}$$

Thus the expected area of S_{abcd} is:

$$\bar{S}_{abcd} = 2\pi(2R)^2 - \bar{S}_{aecf} = 4.8435\pi R^2 \tag{3}$$

Therefore we can derive the probability of collision $p_{intf} = \frac{S_{abcd}}{A}$ where A is the region of area that the nodes are employed. To sum up we now have the underlying constraints:

$$\begin{cases} N_{effective} = p_1 p_2 N p_{intf} = N(\frac{4.8435\pi R^2}{Ar(r+1)}) \\ M^{d+1} \geq 4|S|N_{effective} \\ |S| > (N_{effective} - 1)d \\ |S| devides M - 1) \\ M|S| \ minimized \end{cases} \tag{4}$$

To get an approximate value of required orthogonal channel and sub-slot times we consider a simple tracking scenario. For a simple tracking scenario, the density of sensor nodes should be large enough so that at any given time and location in the two-dimensional area under its coverage, at least three sensors be simultaneously able to sense the moving object with their normal sensing range. Since the number of sensors is large, the distribution of the number of nodes in any given area A is Poisson with rate λA in which λ $nodes/m^2$ is the nodes density. Therefore to have at least three sensors with their normal sensing range in any given point with probability 0.99, by substituting a value for R as the sensing range the required node density λ can be calculated by solving (5), given below:

$$0.99 = \sum_{i=0}^{\infty} (e^{-\lambda\pi R^2} \frac{(\lambda\pi R^2)^i}{i!}). \tag{5}$$

Now we can sum up the following parameters: We randomly place $N = 400$ nodes on a square area of $1,000$ m by $1,000$ m. The radio transceiver range is 100 m. In this case we have at least three sensors within their normal sensing range in any given point with probability 0.9997. Using these defined parameters we have plotted the table I. Using our proposed method we will be able to service $p_1 = \frac{k}{r+1}$ of the links. As can be seen from the table I, with the increase in r the required numbers of orthogonal channels and sub-time slots decrease but this needs an increase in the numbers of keys per node k to service the same number of links. Therefore, there is a tradeoff between the required number of orthogonal channels and sub-time slots and the probability of a link to be serviced. Also, with the increase in the degree of the employed polynomial d the required number of orthogonal channels and sub-time slots reduces but this will lead to more computation complexity in the proposed algorithm.

4 Conclusion

We have addressed the problem of countering the control channel jamming against internal adversaries in wireless sensor networks using key distribution. Using combinotrial approach introduces the degree of the combinatorial design (r) as a controllable parameter which enables the designer to avoid unwanted channel interference and arrive at the prameters that are consistent with the

Table 1. The Required Number of Orthogonal channels and Sub-time Slots for different range of r for $A = 10^6\ m^2$, $N = 400$ and $R = 100\ m$

M	$\|S\|$	r	d	$\frac{1}{r+1}$
10	18	3	2	1/4
6	10	4	2	1/5
9	8	4	3	1/5
6	10	4	4	1/5
6	5	5	4	1/6
3	2	7	4	1/8

needs of WSN and ad-hoc networks . This method can be considered as a special case of the ALOHA random access schemes which uses the common established keys as its seeds to generate the pattern of transmission. Through numerical analysis it is shown that the required resources of this scheme are consistent with the needs of wireless ad hoc networks. It is also demonstrated that assigning channels using this method introduces a tradeoff between required parameters of the scheme such as number of orthogonal channels, sub-time slots, degree of combinatorial design and degree of the used polynomial in the hash function. However this scheme is not a communication efficient scheme but is the first scheme that can fulfill the security and reliability tasks of secure control channel allocation schemes based on key distribution scenarios.

References

1. Fazel, K., Kaiser, S.: Multi-Carrier and Spread Spectrum Systems. Wiley, Chichester (2003)
2. Anderson, R.: Security Engineering: A Guide to Building Dependable Distributed Systems. Wiley, Chichester (2001)
3. Li, M., Koutsopoulos, Poovendran, R.: Optimal jamming attacks and network defense policies in wireless sensor networks. In: 26th IEEE International Conference on Computer Communications, pp. 1307–1315. IEEE Press, USA (2007)
4. Chan, A., Liu, X., Thapa, B.: Control Channel Jamming: Resilience and Identification of Traitors. In: ISIT (2007)
5. Tague, P., Li, M., Poovendran, R.: Mitigation of Control Channel Jamming unde Node Capture Attacks. IEEE Transactions on Mobile Computing 8 (2009)
6. Lazos, L., Liu, S., Krunz, M.: Mitigating control-channel jamming attacks in multichannel ad hoc networks. In: 2th ACM Conference on Wireless Network Security (2009)
7. Campete, S.A., Yener, B.: Combinatorial design of key distribution mechanisms for wireless sensor networks. IEEE/ACM Trans. Network. 15, 346–358 (2008)
8. Lee, J., Stinson, D.R.: A combinatorial approach to key predistribution for distributed sensor networks. In: IEEE Wireless Communications and Networking Conference. IEEE Press, Los Alamitos (2001)
9. Cametpe, S.A., Yener, B., Yung, M.: Expander graph based key distribution mechanisms in wireless sensor networks. In: IEEE Int. Conf. on Commun. IEEE Press, Los Alamitos (2006)

10. Bhandari, V., Vaidya, H.: Secure capacity of multi-hop wireless networks with random key pre-distribution. In: Proc. IEEE INFOCOM Workshops, Workshop on Mission Critical Networking (MCN). IEEE Press, Los Alamitos (2008)
11. Noubir, G.: Collision-Free One-Way Communication Using Reed-Solomon Codes. In: IEEE International Symposium on Information Theory and Applications. IEEE Press, New Mexico (1998)

A Proxy Signature Scheme Based on Coding Theory

Hoda Jannati and Abolfazl Falahati

Department of Electrical Engineering (DCCS Lab)
Iran University of Science and Technology, Tehran, Iran
{hodajannati,afalahati}@iust.ac.ir

Abstract. Proxy signature helps the proxy signer to sign messages on behalf of the original signer. This signature is used when the original signer is not available to sign a specific document. In this paper, we introduce a new proxy signature scheme based on Stern's identification scheme whose security depends on syndrome decoding problem. The proposed scheme is the first code-based proxy signature and can be used in a quantum computer. In this scheme, the operations to perform are linear and very simple thus the signature is performed quickly and can be implemented using smart card in a quite efficient way. The proposed scheme also satisfies unforgeability, undeniability, non-transferability and distinguishability properties which are the security requirements for a proxy signature.

Keywords: code-based signature, proxy signature, Stern's identification, syndrome decoding.

1 Introduction

In 1978, McEliece introduced a public key cryptosystem based on a well-known error-correcting codes problem [9]. After 30 years of research, error-correcting codes problem is still exponential. Besides this, the error-correcting codes problem has three advantages as: (i) In a quantum computer, it can be used in place of the number-theory based problems like the factorization and the discrete logarithm problems, (ii) It is based on linear and very simple operations and (iii) The secret key, for error-correcting codes problem based protocols, is smaller than that of the other protocols for the same confidence level.

Recently kinds of the code-based signature schemes have been proposed. The signature scheme proposed in [3,5], the identification and signature scheme proposed by Stern in [11] and the threshold ring signature scheme proposed in [10] are based on error-correcting codes problem usually referred as the Syndrome Decoding Problem. Thus it is interesting to propose some other cryptosystems such as proxy signature, group signature and blind signature schemes based on syndrome decoding problem.

In proxy signature scheme, a user, which is called original signer, delegates its signing capability to another user, which is called proxy signer. In 1996, the

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 282–290, 2010.

concept of the proxy signatures, which are used when the original signer is not available to sign a specific document, has been introduced by Mambo et al. in [8]. Afterwards, a number of proxy signature schemes have been proposed. They are based on intractability of discrete logarithm problem [6,7], factorization problem [12] and bilinear pairing [13]. Proxy signatures have become very important and popular tools in various e-services such as e-commerce, e-voting and e-health; straightforwardly there is a growing interest in concrete implementations.

Our Contribution. The identification scheme constructed by Stern [11] is one of today's most complete identification schemes based on syndrome decoding problem [2]. In this paper, we propose a proxy signature scheme based on Stern's identification scheme whose security depends on syndrome decoding problem. The proposed proxy signature scheme is the first code-based proxy signature scheme. In this scheme, the operations to perform are linear and very simple (the multiplication of a vector by a matrix or the permutation on words) therefore the signature is performed quickly and can be implemented in both hardware and software in a quite easy and efficient way. The overall complexity of the proposed scheme is a few more than that of Stern's scheme. In this paper, it is also shown the scheme satisfies the security properties a proxy signature scheme such as unforgeability, undeniability, non-transferability and distinguishability properties.

Organization. The remainder of this paper is organized as follows: in Section 2, the properties of a proxy signature scheme are described. In Section 3, Stern's identification scheme is introduced. Then a proxy signature scheme which is provided by Stern's scheme is proposed in Section 4 and discussed specifically on the security and the performance of the proposed scheme in Section 5. Finally, we summarize our research in Section 6.

2 Overview of Proxy Signature

In this section, the basic definitions and the security properties of a proxy signature are explained [8]. There are three main types of players in this signature: (i) original signer, (ii) proxy signer and (iii) verifier. The basic version of a proxy signature scheme consists of four phase as follows:

1. Key generation; the original signer and the proxy signer enter the system and select their secret and public keys.
2. Proxy generation; the original signer generates a proxy key using its secret key for the proxy signer.
3. Signing; the proxy signer generates a valid proxy signature rather than original signer using the proxy key for the verifier.
4. Verifying; the verifier begins a challenge/response interaction with the proxy signer in order to verify the validity of the signature and the proxy signer.

In a proxy signature scheme only the original signer can generate a valid proxy key; even the proxy signer cannot generate another proxy key based on the

received one, and only the proxy signer can generate a valid proxy signature for the verifier; even the original signer cannot do this (strong unforgeability). The original signer cannot repudiate a previous delegation of its signing capability to a proxy signer (strong undeniability). The proxy signer cannot transfer the proxy key given by the original signer to others (non-transferability). Also, a proxy signature scheme must be distinguishable from a normal signature scheme; Means, a verifier should have the ability to recognize that the message has been signed by the proxy signer or the original signer (distinguishability).

3 Overview of Stern's Identification

3.1 Difficult Problem in Coding Theory

Let us recall a linear binary code C of length n and dimension k is a vector subspace of dimension k of $GF(2)^n$. The weight of an element x of $GF(2)^n$ is the number of non zero coordinates of x. A generator matrix G, $(n \times k)$, of a code is a generator basis of a code, the dual of code C is defined by $C^{dual} = \{c' \in GF(2)^n | c \cdot c' = 0, \forall c \in C\}$. Usually a generator matrix of the dual of a code C is called the parity check matrix of C and denoted by $(n-k \times n)$ matrix H. Remark that $\forall c \in C \leftrightarrow H \times c^T = 0$. For $x \in GF(2)^n$ the value $H \times x^T$ has a length of $(n-k)$bits and is called the syndrome of x for H [11]. The usual hard problem considered in coding theory is the following Syndrome Decoding (SD) problem:

Problem: Syndrome Decoding (SD).
Instance: A vector $f \in GF(2)^{(n-k)}$, a $(n-k \times n)$ random matrix H over $GF(2)$, and an integer t.
Question: Is there a vector $s \in GF(2)^n$ of weight $\leq t$ such that $H \times s^T = f$?

In 1978, syndrome decoding problem was proven NP-complete [1]. Stern proposed his identification scheme based on this problem [11].

3.2 Notations

In order to describe stern's scheme, we will use the following notations:

- x^T : transpose of vector x,
- n, k, t : three integers such that $t < k < n$,
- H : a $(n-k \times n)$ public matrix over $GF(2)$,
- s_u : the secret key of user U where $s_u \in GF(2)^n$,
- H_u : the public key of user U where $H_u = H \times {s_u}^T$ and $H_u \in GF(2)^{(n-k)}$,
- $h(\cdot)$: a secure one-way hash function with output length of l-bits,
- $\|$: the operation of concatenation.

3.3 Stern's Identification Scheme

In 1993, Stern proposed an identification scheme based on coding theory. In this scheme, a user U identifies itself to another user, which is called the verifier V [11]. This scheme is described in four steps:

– *Commitment Step*: U randomly chooses $y \in GF(2)^n$ and a permutation σ defined over $GF(2)^n$, computes (1) to (3) and sends (c_1, c_2, c_3) to V.

$$c_1 = h(\sigma \| (H \times y^T)) \tag{1}$$

$$c_2 = h(\sigma(y)) \tag{2}$$

$$c_3 = h(\sigma(y \oplus s_u)) \tag{3}$$

– *Challenge Step*: V after receiving (c_1, c_2, c_3) sends $b \in \{0, 1, 2\}$ to U.

– *Answer Step:*

- If U received $b = 0$: U sends y and σ to V.
- If U received $b = 1$: U sends $(y \oplus s_u)$ and σ to V.
- If U received $b = 2$: U sends $\sigma(y)$ and $\sigma(s_u)$ to V.

– *Verification Step:*

- If $b = 0$: V checks the validity of $c_1 \stackrel{?}{=} h(\sigma \| (H \times y^T))$ and $c_2 \stackrel{?}{=} h(\sigma(y))$.
- If $b = 1$: V checks the validity of $c_1 \stackrel{?}{=} h(\sigma \| (H \times (y \oplus s_u)^T \times H_u))$ and $c_3 \stackrel{?}{=} h(\sigma(y \oplus s_u))$.
- If $b = 2$: V checks the validity of $c_2 \stackrel{?}{=} h(\sigma(y))$, $c_3 \stackrel{?}{=} h(\sigma(y) \oplus \sigma(s_u))$ and that the weight of $\sigma(s_u)$ is t.

If so, V accepts U as a valid user; otherwise, U is an invalid user.

As proved in [11], to get a security level of β, the verifier must iterate the protocol with user U a number of times i such that $(2/3)^i < \beta$. Also, by using the Fiat-Shamir Paradigm [4], this identification scheme can be converted into a signature scheme.

4 The Proposed Proxy Signature Scheme

In this section, we propose a proxy signature scheme, which is based on Stern's identification scheme and whose security depends on syndrome decoding problem. There are three players in the scheme: the original signer O, the proxy signer P and the verifier V and there are four phases in the scheme:

1. Key generation:
   When initializing the system:

– A central party has to choose appropriate values for the parameters n, k and t. Also, it selects a $(n-k \times n)$ matrix H over $GF(2)$. Then it publishes these values as public parameters. The security of the system depends on the choice of these parameters.

– O chooses randomly $s_o \in GF(2)^n$ where the weight of s_o is t and sets $H_o = H \times {s_o}^T$. Then O publishes H_o as the public key and stores s_o as a secret key.

– P chooses randomly $s_p \in GF(2)^n$ where the weight of s_p is t and sets $H_p = H \times {s_p}^T$. Then P publishes H_p as the public key and stores s_p as a secret key.

2. Proxy generation:
   – O chooses randomly a $(n \times l)$ matrix Y with weight w, sets a parameter m_w which contains condition and duration of proxy, computes (4) to (6) and sends (s'_{pr}, Y_o, m_w) to P.

$$Y_o = H \times Y \tag{4}$$

$$m = h(H_o \| H_p \| Y_o \| m_w) \tag{5}$$

$$s'_{pr} = s_o \oplus (Y \times m^T)^T \tag{6}$$

   – P after receiving proxy key s'_{pr}, computes the secret proxy key as

$$s_{pr} = s'_{pr} \oplus s_p \tag{7}$$

   then checks its correctness using (8).

$$H_{pr} = H \times s_{pr}^T \stackrel{?}{=} H_o \oplus H_p \oplus (Y_o \times h(H_o \| H_p \| Y_o \| m_w)^T) \tag{8}$$

   If (s_{pr}, Y_o, m_w) satisfies (8), P accepts s_{pr} as a secret proxy key and publishes H_{pr} as a public proxy key on behalf of O.

3. Signing:
   In this phase, the proxy signer P signs for the verifier V on behalf of the original signer O using Stern's scheme. The details are as follow:

   – *Commitment Step*: P randomly chooses $z \in GF(2)^n$ and a permutation σ defined over $GF(2)^n$, computes (9) to (11) and sends $(c_1, c_2, c_3, Y_o, m_w)$ to V.

$$c_1 = h(\sigma \| (H \times z^T)) \tag{9}$$

$$c_2 = h(\sigma(z)) \tag{10}$$

$$c_3 = h(\sigma(z \oplus s_{pr})) \tag{11}$$

– *Challenge Step*: V after receiving $(c_1, c_2, c_3, Y_o, m_w)$ checks the validity of (12).

$$H_{pr} \stackrel{?}{=} H_o \oplus H_p \oplus (Y_o \times h(H_o \| H_p \| Y_o \| m_w)^T) \tag{12}$$

If so, V accepts P as a valid proxy signer on behalf of O and then sends $b \in \{0, 1, 2\}$ to P.

– *Answer Step:*
- If P received $b = 0$: P sends z and σ to V.
- If P received $b = 1$: P sends $(z \oplus s_{pr})$ and σ to V.
- If P received $b = 2$: P sends $\sigma(z)$ and $\sigma(s_{pr})$ to V.

4. Verifying:
– If $b = 0$: V checks the validity of (13) and (14).

$$c_1 \stackrel{?}{=} h(\sigma \| (H \times z^T)) \tag{13}$$

$$c_2 \stackrel{?}{=} h(\sigma(z)) \tag{14}$$

– If $b = 1$: V checks the validity of (15) and (16).

$$c_1 \stackrel{?}{=} h(\sigma \| (H \times (z \oplus s_{pr})^T \times H_{pr})) \tag{15}$$

$$c_3 \stackrel{?}{=} h(\sigma(z \oplus s_{pr})) \tag{16}$$

– If $b = 2$: V checks the validity of (17) and (18) and that the weight of $\sigma(s_p r)$ is $\leq 2t + w$.

$$c_2 \stackrel{?}{=} h(\sigma(z)) \tag{17}$$

$$c_3 \stackrel{?}{=} h(\sigma(z) \oplus \sigma(s_{pr})) \tag{18}$$

If so, V accepts it as a valid signature on behalf of O; otherwise, it is an invalid signature.

5 Security and Performance

5.1 Security Consideration of the Proposed Proxy Signature Scheme

In this section, it is shown that the proposed proxy signature scheme satisfies all the security requirements defined in section 2.

– *Strong Unforgeability*
a) "Nobody can generate a valid proxy key except the original signer." Suppose that one, e.g., the proxy signer, tries to forge a proxy key. From (6), we know that the proxy key includes the secret key s_o of the original signer which is clearly unknown to the others. Therefore, the secret key

of the original signer must be extracted from (6). Since Y is a random $(n \times l)$ matrix, it is difficult to derive the value of s_o. To find Y from Y_o, it is needed to solve the syndrome decoding problem or guess it with at most a probability $(n \times l)^{-1}$. Therefore, the security of the proxy generation phase of the proposed scheme is linked to syndrome decoding problem.

b) "Nobody can create a valid proxy signature except the proxy signer." In the proposed scheme, signing phase is based on Stern's identification scheme. Thus, a forger can sign in place of the proxy signer if and only if she can break Stern's identification scheme or have s_{pr}. Note that if someone tries to find s_{pr} from H_{pr} again, it is needed to solve the syndrome decoding problem. The original signer has (s'_{pr}, Y_o), but cannot obtain s_{pr} because $s_{pr} = s'_{pr} \oplus s_p$ and s_p is unknown to her. The original signer can obtain it properly by guessing with at most a probability of n^{-1}.

– *Strong Undeniability*
"O cannot repudiate the delegation of its signing capability to a proxy signer." For executing a correct proxy signature, (12) must be satisfied. This equation includes the public key H_o of the original signer. So, this equation is correct if s'_{pr} is generated based on the secret key of O. It has been proved in part a) of *Strong Unforgeability* that only O can generate proxy key s'_{pr}. Therefore, the original signer O cannot repudiate generating of the proxy key s'_{pr} for the proxy signer.

– *Non-transferability*
P cannot transfer the proxy key given by the original signer to others because we inserted some information of the proxy signer in the proxy generation phase. Suppose that P is going to transfer a proxy key generated using H_p and m_w, $s'_{pr} = s_o \oplus (Y \times h(H_o \| H_p \| Y_o \| m_w)^T)$ to other proxy signer which has H_{p2} and s_{p2} as its public and secret keys respectively. Thus, he has to modify the proxy key in order to show that it has been generated using H_{p2} and m_{w2}. However, P do not know s_o and Y, Therefore for correcting (12), P must use the same s'_{pr}. Therefore H_{pr2} is computed using (19).

$$H_{pr2} = H \times {s_{pr2}}^T = H \times (s_{p2} \oplus s'_{pr})^T = H_o \oplus H_{p2} \oplus (Y_o \times h(H_o \| H_p \| Y_o \| m_w)^T) \tag{19}$$

However, according to (12), we should have:

$$H_{pr2} = H_o \oplus H_{p2} \oplus (Y'_o \times h(H_o \| H_{p2} \| Y'_o \| m_{w2})^T) \tag{20}$$

Therefore, to find Y'_o correctly, (19) and (20) must be the same:

$$H_o \oplus H_{p2} \oplus (Y'_o \times h(H_o \| H_{p2} \| Y'_o \| m_{w2})^T) = H_o \oplus H_{p2} \oplus (Y_o \times h(H_o \| H_p \| Y_o \| m_w)^T) \tag{21}$$

$$Y'_o \times h(H_o \| H_{p2} \| Y'_o \| m_{w2})^T = Y_o \times h(H_o \| H_p \| Y_o \| m_w)^T \tag{22}$$

Thus, Y'_o must be found to satisfy (22). However, it is computationally infeasible, since Y'_o contributes to the hash padding too.

– *Distinguishability*
Obviously, the verifier can easily distinguish the original signer's signature from proxy signer's because they use different keys, i.e., (s_o, H_o) for O and (s_{pr}, H_{pr}) for P, and they don't access to each other's secret key.

5.2 Performance of the Proposed Scheme

A quick analysis of the proposed scheme shows that the different phases are composed of actually four main operators: the multiplication of a vector by a matrix, the action of a hash function, the generation and the action of a random permutation on words. Therefore, the operations to perform are linear and very simple.

In this scheme, the signature has been based on Stern's identification scheme. In [2], authors have presented the implementation on smartcard of the Stern's identification and signature scheme. They have given a secure scheme against side channel attacks and obtain the identification in around 5.5 seconds and the signature in around 22 seconds for a security of 2^{85}. Table 1 summarizes the proposed proxy signature cost compared with stern's scheme cost.

Table 1. The proposed proxy signature cost compare to Stern's scheme cost

	Proxy Generation Cost	Signing Cost	Verifying Cost	Total Cost
	hash,$\oplus$,$\times$	hash,$\oplus$,$\times$	hash,$\oplus$,$\times$	hash,$\oplus$,$\times$
Stern's scheme	-	$3i$,i,i	$2i$,i,$2i$	$5i$,$2i$,$3i$
The proposed scheme	2,4,4	$4i$,$3i$,$2i$	$2i$,i,$2i$	$6i+2$,$4i+4$,$4i+4$

Proxy generation phase cost is negligible against to quasi-cyclic construction of signing and verifying phases, therefore the higher costs in the proposed proxy signature scheme are signing and verifying phases. Thus, according to Table 1 and the quasi-cyclic construction of signing and verifying phases, we can confirm the total cost of our signature is a few more than that of Stern's scheme. Therefore, this scheme can be implemented on smartcard in a quite efficient way, the same as stern's scheme.

6 Conclusion

We proposed a proxy signature scheme based on Stern's scheme whose security relies on syndrome decoding problem. In the proposed proxy signature scheme, only the original signer can generate a valid proxy key and only the proxy signer can generate a valid proxy signature. Also the possession of a proxy key which has been delegated to a proxy signer cannot be repudiated by the original signer. In the proposed scheme, the signature is performed quickly because the operations to perform are linear and very simple (the multiplication of a vector by a matrix

or the permutation on words), therefore this scheme can be implemented in hardware as well as software in a quite easy and efficient way. Also, the overall complexity of the scheme is a few more than that of Stern's scheme. Besides these properties the scheme present the first code-based proxy signature scheme and can open a new area of research since the proxy signature schemes today are employed in many different IT areas such as e-voting, e-health and e-commerce protocols.

References

1. Berlekamp, E., McEliece, R., Tilborg, H.: On the Inherent Intractability of Certain Coding Problems. IEEE Transactions on Information Theory 24, 384–386 (1978)
2. Cayrel, P.-L., Gaborit, P., Prouff, E.: Secure Implementation of the Stern Authentication and Signature Schemes for Low Resource Devices. In: Grimaud, G., Standaert, F.-X. (eds.) CARDIS 2008. LNCS, vol. 5189, pp. 191–205. Springer, Heidelberg (2008)
3. Courtois, N., Finiasz, M., Sendrier, N.: How to achieve a MCEliece-based Digital Signature Scheme. In: Boyd, C. (ed.) ASIACRYPT 2001. LNCS, vol. 2248, pp. 157–174. Springer, Heidelberg (2001)
4. Fiat, A., Shamir, A.: How to Prove Yourself: Practical Solutions to Identification and Signature Problems. In: Odlyzko, A.M. (ed.) CRYPTO 1986. LNCS, vol. 263, pp. 186–194. Springer, Heidelberg (1987)
5. Hamdi, O., Harari, S., Bouallegue, A.: Secure and Fast Digital Signatures using BCH Codes. International Journal of Computer Science and Network Security 6(10), 220–226 (2006)
6. Jannati, H., Salmasizadeh, M., Mohajeri, J.: New Proxy Signature, Proxy Blind Signature and Blind Proxy Signature Based on the Okamoto Signature. In: International Conference on Security and Management, pp. 238–242. SAM, Las Vegas (2008)
7. Lu, R., Cao, Z., Zhou, Y.: Proxy Blind Multi-Signature Scheme Without a Secure Channel. Journal of Applied Mathematics and Computation 164(1), 179–187 (2005)
8. Mambo, M., Usuda, K., Okamoto, E.: Proxy Signatures for Delegating Signing Operation. In: Conference on Computer and Communications Security - CCS 1996, pp. 48–57. ACM Press, New York (1996)
9. McEliece, R.-J.: A Public-Key Cryptosystem Based on Algebraic Coding Theory. DSN Prog. Rep., Jet Propulsion Laboratory, California Inst. Technol., Pasadena, CA, pp. 114-116 (January 1978)
10. Melchor, C.-A., Cayrel, P.-L., Gaborit, P.: A New Efficient Threshold Ring Signature Scheme based on Coding Theory. In: Buchmann, J., Ding, J. (eds.) PQCrypto 2008. LNCS, vol. 5299, pp. 1–16. Springer, Heidelberg (2008)
11. Stern, J.: A New Identification Scheme Based on Syndrome Decoding. In: Stinson, D.R. (ed.) CRYPTO 1993. LNCS, vol. 773, pp. 13–21. Springer, Heidelberg (1994)
12. Wang, G., Bao, F., Zhou, J., Deng, R.-H.: Comments on a Practical (t,n) Threshold Proxy Signature Scheme Based on the RSA Cryptosystem. IEEE Transaction on Knowledge and data Engineering 16(10), 1309–1311 (2004)
13. Zhang, F., Safavi-Naini, R., Susilo, W.: An Efficient Signature Scheme from Bilinear Pairings and its Applications. In: Bao, F., Deng, R., Zhou, J. (eds.) PKC 2004. LNCS, vol. 2947, pp. 277–290. Springer, Heidelberg (2004)

Partially Key Distribution with Public Key Cryptosystem Based on Error Control Codes

Saeed Ebadi Tavallaei and Abolfazl Falahati

Department of Secure Communication (DCCS Lab),
Iran University of Science and Technology, Narmak, Tehran
sebadi@ee.iust.ac.ir,afalahati@iust.ac.ir
http://www.iust.ac.ir

Abstract. Due to the low level of security in public key cryptosystems based on number theory, fundamental difficulties such as "key escrow" in Public Key Infrastructure (PKI) and a secure channel in ID-based cryptography, a new key distribution cryptosystem based on Error Control Codes (ECC) is proposed . This idea is done by some modification on McEliece cryptosystem. The security of ECC cryptosystem obtains from the NP-Completeness of block codes decoding. The capability of generating public keys with variable lengths which is suitable for different applications will be provided by using ECC. It seems that usage of these cryptosystems because of decreasing in the security of cryptosystems based on number theory and increasing the lengths of their keys would be unavoidable in future.

Keywords: Key Distribution, Error Control Codes (ECC), Public Key Cryptosystem (PKC), Public Key Infrastructure (PKI), NP-complete problem.

1 Introduction

Basically the key-security plays an important role in any cryptosystem.Though key distribution, being one of the most significant factors that affect the security of a cryptosystems is more important than choosing the encryption algorithm to reach the security. Key distribution cryptosystems are divided into two main groups: Public Key Infrastructure (PKI) and Identity-based cryptosystems [1]. In a traditional Public Key Infrastructure after generating the keys, the Certificate Authority (CA) certifies the user's public key. In a case of using a public key, each participant has to primarily verify the corresponding certificate. The public keys are verified by Certificate Authority and are given to each user. As a consequence there is a need for key revocation. These processes require much more processing time as well as the capacity. Id-based cryptography that is recommended by Shamir in 1984, considerably simplifies the key management problem. In the Id-based cryptography the user's public key is directly derived from user's identity information i.e. his/her name, email, etc. Then the corresponding user's private key, assuring with trusted third party named Key Generation Center (KGC) is

S. Tenreiro de Magalhães, H. Jahankhani, and A.G. Hessami (Eds.): ICGS3 2010, CCIS 92, pp. 291–299, 2010.

generated and given to the users through a secure channel. As there is no need for key producing and revoking, Id-based cryptography compared with Public Key Infrastructure (PKI) has many advantages in key management. In Id-based cryptography the user sends a secure message to the receiver using the receiver's identity information, even before the receiver's gets his own private key from KGC [2].

An inherent problem of Id-based cryptography is the key escrow problem, i.e. user's private key is known to the KGC's. This enables the KGCs to decrypt any ciphertext and forge signature for any message, leaving neither the user privacy nor the system authenticity [1]. It also requires a secure channel between users and the KGC to deliver private keys. Because of these inherent difficulties, Id-based cryptography is considered to be suitable only for small private networks with lower security requirements. This paper deals with a certificate-based key distribution using Error Control Codes (ECC). This idea done by developing coding-based cryptosystems such as Identification [3], digital signature [4], hash functions [5] and pseudo-random generator [6] cryptosystems. The security of these cryptosystems is obtained from the inherent intractability decoding of block codes [7]. One of the security policies for increasing the level of security in key distribution is using several KGCs. Doing key generation and key issuing by several KGCs causes to divide the work load of each KGC and also reduces successful attacks on these KGCs. The proposed method uses the ECC cryptosystem, which was Introduced by McEliece in 1978 to develop a public key cryptosystem [8]. With some basic changes in McEliece cryptosystem we improve it for key distribution. In addition for increasing the security, private keys are generated by several KGCs, forming a partial key distribution which prevents common attacks on KGCs. Instead of sending the key itself, in the partial key distribution, its parameters are sending to the user. In fact this enables the user to generate his/her private keys without being informed by other KGCs. In the following we discuss the original model of McEliece and the proposed model respectively.

2 McEliece Cryptosystem

The McEliece public key cryptosystem is based on a block code $C(n, k, t)$, where k, n are the length of the input and output bits respectively and t is the error correction capability of the code. In the McEliece cryptosystem, the public key is generated by the secret keys as: $G_{pub} = S.G.P$ where the matrices $G_{[k\times n]}$, $P_{[n\times n]}$ and $S_{[k\times k]}$ are the systematic generator matrix, permutation matrix and non-singular scrambler matrix respectively. The encryption of a message m is performed by adding a random error vector e, of weight $w_H(e) = t$ as:

$$c = m.G_{pub} + e = m.S.G.P + e \tag{1}$$

The decryption of ciphertext c can be performed as follows: By multiplying c by P^{-1} we have $m.S.G.P.P^{-1} + e.P^{-1}$ With a fast-decoding algorithm which is

a high speed decoding algorithm such as Goppa codes [9] [10] we will have the $m.S$

Finally by multiplying S^{-1} we obtain the message m. A simple method to modify the McEliece cryptosystem is by using a hiding matrix as [11]:

$$G_{pub} = S.G.P + X \tag{2}$$

where X is a specific matrix of rank one. The random matrix X is chosen as an extra secret key and is added to the original public key to make a new modified public key. Thus, any visible structure of the public key will be hidden. On the other hand, the hiding matrix should have as many free parameters as possible. This problem can be solved at the expense of having a lower error-correcting capacity needed by the decoder from t to $t-1$. Let m be an information vector to be encrypted. A message is then encrypted as:

$$c = m.G_{pub} + e = m.S.G.P + e \tag{3}$$

$where w_H(e) \leq t-1$

3 The Recommended Scheme

The proposed pubic key structure is similar to (3) and is considered as:

$$G_{pub} = S(G + X_1).P + X_2 \tag{4}$$

where the private keys (i.e. S, G, P and X_1) are generated by some separated KGCs as shown in Fig. 1 .

Key Generation done through these stages:
1- parameters requesting
2- receiving (n, k)
3- sending (n, k) and requesting the other parameters
4- receiving $p(x)$

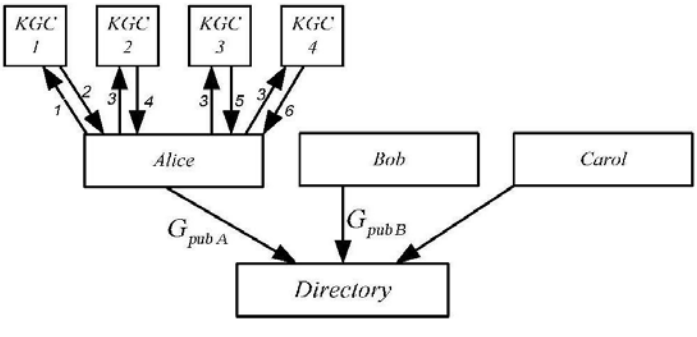


Fig. 1. Key Distribution scheme

5- receiving S; $p(x)$ is a primitive polynomial
6- receiving X

$$G_{pub_A} = S(G + X_1).P + X_2$$
$$G_{pub_B} = S(G + X_2).P + X_2$$

3.1 First Stage: Requesting the Parameters of the Private Keys

In this stage to get a public key, each user with their personal identity (like email) refers to one of the KGCs and requests the parameters of generator matrix (i.e. G). In order to increase the security level of the system we generate the G parameters by different KGCs i.e. each KGC sends only one parameter of the code (i.e. n, k and *primitivepolynomial*) to the user. After obtaining the parameter n of the code, there are some choices for choosing the parameter k, which can be selected by one of the KGCs based on security policy of the cryptosystem. Then the user refers to another KGC which is completely separated from the previous ones, and after sending his personal identity and (n, k) of the produced code, requests a primitive polynomial $p(x)$, under the certain Galois field. This KGC select a primitive polynomial and sends it to the desired user securely. Now the user knows completely the characteristics of the certain block code like $C(n, k)$. Note that the decoding of a block codes without knowing the characteristics of the code is an NP-complete problem [7].

3.2 Second Stage: Producing the Private Keys

As all the parameters needed for a code are obtained, the user is able to generate the minimal polynomial of the code to generate the systematic form of generator matrix (i.e. G) where this is the main part of his private keys. To obtain other private keys (i.e. S , P and X), the user refer to the other KGCs. After introducing himself and sending (n, k) he securely receives other private keys. Of course with considering the security policy which may be and no needs for more authenticity for the key, he can produce these private keys personally.

3.3 Third Stage: Producing the Pubic Key

The user computes his pubic key as (2) and announces it publicly, so that others can communicate with him securely.

3.4 Forth Stage: Establishing the Connection

To verify the public key the user refers to the Certificate Authority (CA), and receives the hashed public key to encrypt the message m, with G_{pub} and sends the encrypted c as $c = m.G_{pub}$ to the recipient. The owner is the only one who can decrypt the encrypted message c, to derive the original message m.

Then the encryption and decryption process can be done as follow: The encryption of a message m is performed by adding a random error vector e, of weight $w_H(e) = t - 1$ as:

$$c = m.G_{pub} + e = m.S(G + X)P + e \tag{5}$$

The decryption of ciphertext c can be performed by multiplying to P^{-1} :

$$m.S.G.P.P^{-1} + e.P^{-1}$$

$$m.S.G + (m.S.X + e)P^{-1} \tag{6}$$

Now with a fast-decoding algorithm which is a high speed decoding algorithm such as Goppa codes [9][10] we will have the $m.S$

Note that the decoding of a block code without knowing the characteristics of the code is an NP-complete problem [7]. Finally by multiplying $m.S$ by S^{-1} the message m will be obtain. When the users want to authenticate and exchange their keys they can do as shown in Fig. 2

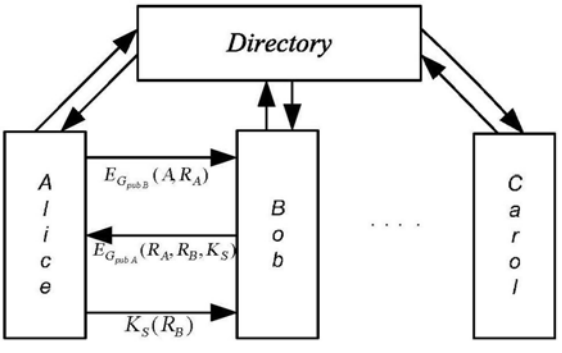


Fig. 2. Authentication and public key exchange

4 The Security of Recommended Cryptosystem

The attack on the proposed cryptosystem can be done in two types: the first type includes the attacks on the cryptosystem algorithm, while the second type includes the conventional attacks on the KGCs. We consider some attacks in Fig. 3

4.1 Type One: Attacks on the Cryptosystem

Although the structure of the generator matrix is hidden by the matrices X and P, but by using the coding based scheme we reach to a randomized cryptosystem [12] in the process of encryption, that depends on the $(n, k, p(x))$ of the code.

Type One: Cryptosystems Algorithm Attacks	1. Guessing S , P and X 2. Exhaustive Codeword comparison 3. Syndrome Decoding 4. Guessing k correct and Independent Coordinate 5. Multiple Encryption of the same message
Type Two: KGC Attacks	1. Reflection attack 2. Man-in-the-Middle

Fig. 3. Types of Attacks

Without knowing these parameters, the problem of decoding will be an NP-complete problem. There are two types of attacks on the cryptosystem algorithm. The first type known as structural attack which aims to finding the structure or a part of the structure of original code from public key. The second one called direct attack or per message attack, tries to decode a given encrypted message [13] [14] and some are explained in brief.

Attack on the weak keys. For all irreducible binary polynomials $g(x)$ of degree $t = 50$ is approximately 2^{44} , Thus the average runtime of the attack on weak key with the parameters $n = 1024$ will be $(2^{44} + 1)O(n^3 + 2^R n^2 \log(n)) \approx 2^{75}$, where $R = \sum_{i=1}^{\infty} \frac{1}{q^i+1}$.

Guessing S, P and X. The matrices S,P and X are essential for the security of the system. A cryptanalyst should recover the original generator matrix G from G_{pub}. Once he has done that, he can use a decryption algorithm to break the code and find the transmitted message. Therefore, a possible attack is to guess S,P and X.

$S_{[k\times k]}$ is invertible, thus the rows of the matrix must be linearly independent. Therefore, the number of invertible binary matrices of size $[k \times k]$ is $\prod_{i=0}^{k-1}(2^k - 2^i)$. The number of possible permutation matrices $P_{[n\times n]}$ is of course $n!$. And the possible matrices $X_{[k\times n]}$ with rank one are $nk!$ [13] [14].

Exhaustive Codewords Comparison. Another brute force approach is to generate all 2^k codewords in the code defined by G_{pub} to find the unique codeword c which is closet to a candidate vector y, thus $c = m.G_{pub}$ and all distance between c and y, satisfies $d(c, y) \leq t$. With a simple Gaussian elimination process one can then retrieve the hidden message m and compares with 2^k positions. The complexity of this attack will be more than a security needs for a cryptosystem.

Syndrome Decoding. At this method the cryptanalyst could compute the parity check matrix H_{pub} corresponding to G_{pub}, using $G_{pub}.(H_{pub})^T = 0$. This

matrix has rank $n-k$. Next, he can compute the syndrome of the transmitted word y , which is

$$y.(H_{pub})^T = (m.G_{pub} + e).(H_{pub})^T = m.G_{pub}.(H_{pub})^T + e.(H_{pub})^T = e.(H_{pub})^T \quad (7)$$

If he finds the right error vector e, he discovers $m.G_{pub}$ and can easily find m by applying Gaussian elimination. Therefore he can generate all possible error vectors of length n and weight less than t, compute $e.(H_{pub})^T$ and compare these with codeword syndrome. The number of error vectors to try is $\sum_{i=0}^{t} \binom{n}{i}$.

Guessing k correct and Independent coordinates. Sometimes attacking the system can select k random positions in the hope that they have'nt any error. If the restriction of G_{pub} to these k positions still has rank k, one can find a candidate m' for the transmitted message m by applying Gaussian elimination. If the rank is not k, it will very likely be close to k; therefore if one picks k correct positions, the Gaussian process will lead to a few possibilities for m. The probability of the k selected positions are correct is about $(1-t/n)^k$, and the Gaussian elimination involves k^3 steps, so the expected work factor is $k^3.(1-t/n)^{-k}$ [13] [14].

4.2 Type Two: Attack on the KGC

As mentioned before, one type of attack, are the attacks on the KGCs like Reflection attack and man-in-the-middle attack for grabbing user's information. Since the KGCs only send the parameters which are to be published later publicly (parameters like n ,k , t or the $coefficients of polynomial$) to the users, the attack to these KGCs are not successful. Because the main processes of producing the private keys are performed by users themselves without the KGCs being informed.

5 The Advantages of Proposed Cryptosystem

1. As mentioned in third of generating of public keys, the structure of generator matrix G is hidden by P , we increase this masking by adding matrix X. Therefore if the KGCs are poisoned, there would be a little chance for the attacker to obtain the parameters of matrix G .
2. Unlike the present cryptosystems, at this scheme only the parameters of the key (not the key itself) stores in the KGCs, therefore the storage capacity of KGCs are extremely decreased. In addition most of the process for key generation is done by users and this reduces the processing time of KGCs.
3. Even, if some of the KGCs are poisoned, the possibility of generating the private keys (i.e. S, G, P and X) are decreased.
4. One of the difficulties of Id-based cryptography is requiring a secure channel at the beginning of key distribution process. In the recommended cryptosystem there is no need for such a secure channel because the attacker would

not understand the process of key distribution even in the case of active attack. Because of these attacks can be detected quickly.
5. The problem of Key Escrow does not exist in the recommended scheme and the court can identify and authenticate the messages.
6. Partially distribution of key parameters imposes extra work on the users but makes it more complicated for the attackers.
7. The most interesting advantageous of using this scheme is the possibility of generating keys with various lengths which is suitable for different security applications (long or short term keys).
8. When a user requests a key from Certificate Authority, the CA signs the user's name and his public key. It is possible for CA to encrypt the keys with his key and sends it to the applicants so that the user ensures the authenticity of the requested public key.
9. Fig. 4 depicts the main characteristics comparison of proposed cryptosystem with RSA. [15] with $[n = 1024, k = 524, e = 17, de = 1 mod((p-1)(q-1))]$.

	RSA	Proposed
Public key size	$log(n) + log(e)$	$k \times (n-k)$
Private key size	$log(n) + log(d)$	$k \times (n-k) + (n-k+1+2log_2 n) + k^2 + log_2 n$
Rate	1	k/n
Encryption cost	$log(e)log(n)$	$(k/2) + (t/n)$
Decryption cost	$log(d)log(n) < log^2(n)/log(e)$	$n + t + k^2/n + 1$

Fig. 4. Comparing some parameters with RSA

6 Conclusion

Due to the importance of Data Integrity and Data Secrecy in communication systems, especially in wireless channels, using the systems which provides both parameters are important. One of these systems is a cryptosystem which is based on Error Control Codes that could be used for both a cryptosystem and an error correction system. Because of the past technological limitations, no sufficient attention were paid to these systems but with a technological mutation (processing and data storing) and advantages of these cryptosystems like NP-Completeness and high encryption/decryption speed, these systems certainly could be the first choice for future decade. Therefore with public and private cryptosystems based on error control codes like digital signature, authentication, fast hash functions and random number generators is introduced a key distribution cryptosystem based on ECC instead of number theory. In addition the proposed cryptosystem has more advantages against traditional key distribution cryptosystems.

Acknowledgements

The authors would like to thank ITRC (Iran telecommunication Research Center) for their invaluable assistance and funding for this research which is a part of UMTS project.

References

1. Gentry, C.: Certificate-based encryption and the certificate revocation problem. In: Biham, E. (ed.) EUROCRYPT 2003. LNCS, vol. 2656, pp. 272–293. Springer, Heidelberg (2003)
2. Shamir, A.: Identity-Based Cryptosystems and Signature Schemes. In: Blakely, G.R., Chaum, D. (eds.) CRYPTO 1984. LNCS, vol. 196, pp. 47–53. Springer, Heidelberg (1985)
3. Stern, J.: A new identification scheme based on syndrome decoding. In: Stinson, D.R. (ed.) CRYPTO 1993. LNCS, vol. 773, pp. 13–21. Springer, Heidelberg (1994)
4. Alabbadi, M., Wicker, S.B.: A digital signature scheme based on linear error-correcting block codes. In: Safavi-Naini, R., Pieprzyk, J.P. (eds.) ASIACRYPT 1994. LNCS, vol. 917, pp. 238–248. Springer, Heidelberg (1995)
5. Augot, D., Finiasz, M., Sendreir, N.: A family of fast syndrome based cryptographic hash functions. In: Dawson, E., Vaudenay, S. (eds.) Mycrypt 2005. LNCS, vol. 3715, pp. 64–83. Springer, Heidelberg (2005)
6. Fischer, J.B., Stern, J.: An efficient pseudo-random generator provably as secure as syndrome decoding. In: Maurer, U.M. (ed.) EUROCRYPT 1996. LNCS, vol. 1070, pp. 245–255. Springer, Heidelberg (1996)
7. Berlekamp, E.R., McEliece, J.R., van Tilborg, H.: On the Inherent Intractability of Certain Coding Problems. IEEE Trans. Information Theory IT-24, 384–386 (1978)
8. McEliece, R.J.: A Public-Key Cryptosystem Based on Algebraic Coding Theory, Deep Space Network Progress Report 42-44, Jet Propulsion Laboratory, California Institute of Technology, pp. 114–116 (1978)
9. Goppa, V.D.: A new class of linear error-correcting codes. Probl Peredach. Information 6(3), 24–30 (1970)
10. McWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes. North-Holland, Amsterdam (1977)
11. Gabidulin, E.M., Paramonov, A.V., Tretjakov, O.V.: Ideals over a Non-Commutative Ring and their Applications in Cryptography. In: Davies, D.W. (ed.) EUROCRYPT 1991. LNCS, vol. 547. Springer, Heidelberg (1991)
12. Engelbert, D., Overbeck, R., Schmidt, A.: A Summary of McEliece-Type Cryptosystems and their Security, TU-Darmstadt, Department of Computer Science, Cryptography and Computer Algebra Group, May 10 (2006) (preprint)
13. van Tilburg, J. : Security-analysis of a class of cryptosystems based on linear error-correcting codes", Technische Universiteit Eindhoven, Dissertation (1994)
14. van Tilborg, H.C.A.: Fundamentals Of Cryptology, A Professional Reference and Interactive Tutorial. The Kluwer International Series In Engineering And Computer Science, Eindhoven University of Technology. Kluwer Academic Publishers, Dordrecht (2002), eBook ISBN: 0-306-47053-5
15. Canteaut, A., Chabaud, F.: A New Algorithm for Finding Minimum-Weight Words in a Linear Code: Application to McEliece's Cryptosystem and to Narrow-Sense BCH Codes of Length 511. IEEE Trans. Information Theory 44(1), 367–378 (1998)

Author Index