2002

Protecting the infrastructure: 3rd Australian information warfare & security conference 2002

William Hutchinson (Ed.)
Edith Cowan University

Edith Cowan University

Copyright Warning

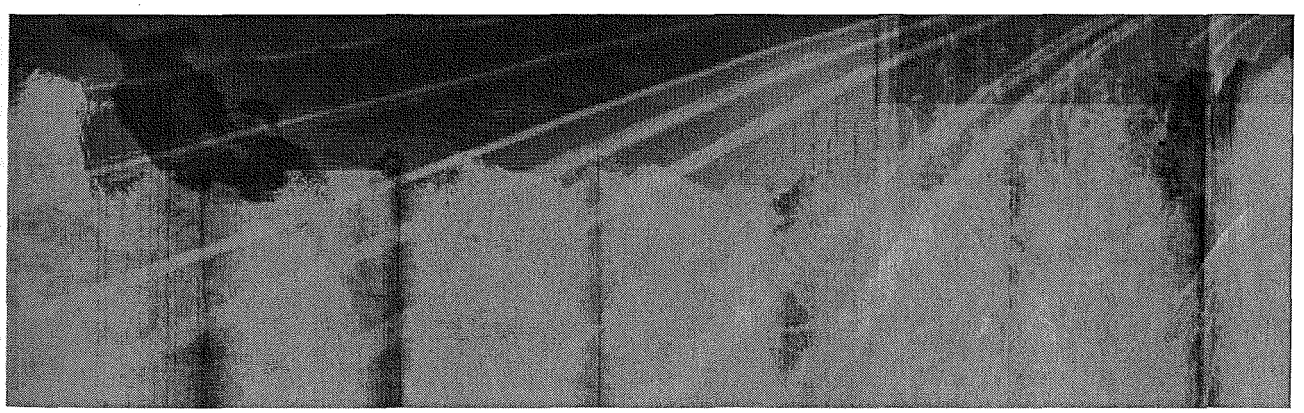
3RD AUSTRALIAN INFORMATION WARFARE & SECURITY CONFERENCE 2002

PERTH, WESTERN AUSTRALIA 28 & 29 NOVEMBER, 2002

"protecting the infrastructure"

P R O C E E D I N G S

EDITORS:
Assoc Professor William Hutchinson
School of Computer & Information Science



ECU EDITH COWAN
school of
Computer & Information Science

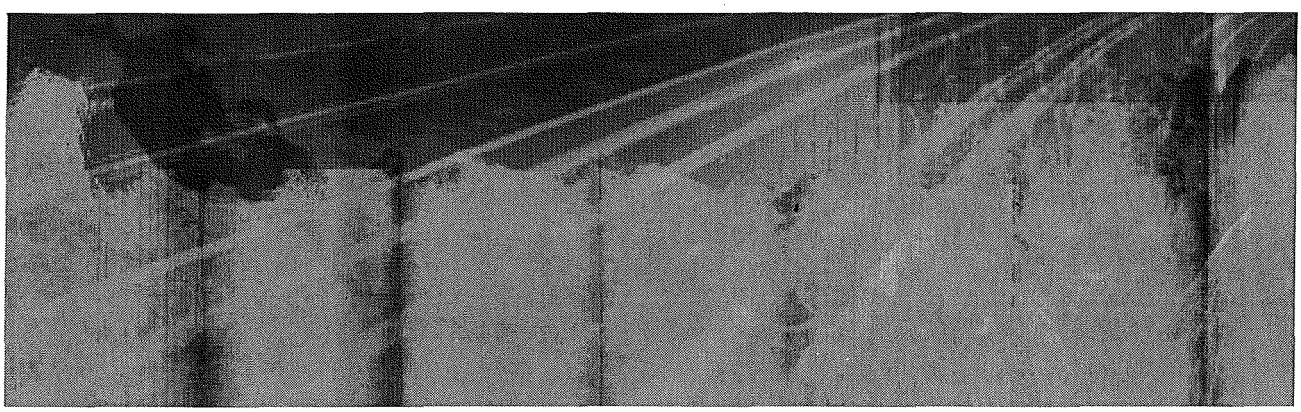
3RD AUSTRALIAN INFORMATION WARFARE & SECURITY CONFERENCE 2002

PERTH, WESTERN AUSTRALIA 28 & 29 NOVEMBER, 2002

"protecting the infrastructure"

PROCEEDINGS AND CD ROM PUBLISHED BY:
We-B Centre | School of Management Information Systems
| School of Computer & Information Sciences
Edith Cowan University
Pearson Street
CHURCHLANDS WESTERN AUSTRALIA 6018

FOR ADDITIONAL COPIES OF THE PROCEEDINGS, CONTACT:
The We-B Centre, School of Management Information Systems
Edith Cowan University
Telephone: +61 8 9273 8607
Facsimile: +61 8 9273 8332
Email: l.davies_moore@ecu.edu.au



ISBN 0-7298-0524-7

All papers that appear in these Proceedings have been subjected to a blind review by two or three anonymous reviewers.

ECU EDITH COWAN UNIVERSITY
school of
Computer & Information Science

FOREWORD

Welcome to the 3rd Australian Information Warfare & Security Conference 2002.

The conference is hosted by the We-B Centre (working with e-business) in the School of Management Information System, the School of Computer & Information Sciences at Edith Cowan University. This year's conference is being held at the Sheraton Perth Hotel in Adelaide Terrace, Perth.
Papers for this conference have been written by a wide range of academics and industry specialists. We have attracted participation from both national and international authors and organisations. The papers cover many topics, all within the field of information warfare and its applications, now and into the future.

The papers have been grouped into six streams:

- Networks
- IWAR Strategy
- Security
- Risk Management
- Social/Education
- Infrastructure

All submitted papers were subjected to a blind review process by two and three anonymous peer reviewers. Reviewers provided both judgement and constructive criticism; several of the papers are even better after the authors followed reviewers' suggestions for improvements. Thank you to all reviewers, for their prompt and professional support.

I would like to thank the Conference Committee and all other people who have helped with the preparation for this conference. In particular I would like to mention:

- Lindsay Davies-Moore, Conference Coordinator, for her tireless efforts at coordinating and managing all the details of the conference
- Professor Janice Burn, Head of School, for her good advice and unfailing support of the concept of the IWAR Conference.

Many people have made valuable contributions to this conference. The most important of these are the authors of all submitted papers. Even those that we could not accept demonstrated good standards of academic and professional effort; I hope that the reviewers' comments will help you to success with future publications. To all of the authors: Thank you.

Welcome to the 3rd Australian Information Warfare & Security Conference 2002.

Enjoy the conference.

ASSOCIATE PROFESSOR WILLIAM HUTCHINSON
Conference Chair
School of Computer & Information Sciences

CONFERENCE COMMITTEE

William Hutchinson	Conference Chair
Craig Valli	Reviewing Committee
Lindsay Davies-Moore	Conference Coordinator
Javen Ang	Multimedia Developer
Shirlee-ann Knight	Web Designer/Master
Prof Janice Burn	Executive Sponsor

CONFERENCE SPONSORS

School of Computer & Information Sciences

We-B Centre & the School of Management Information Systems

Perth Convention Bureau

CONFERENCE REVIEWING COMMITTEE

William Hutchinson
Mathew Warren
Craig Valli
Colin Ash
Jack Davey
Helen Armstrong
Andrew Dawson
Steve Fall
Jill Slay
Phil Dobson
Chris Hu
Mark Williams
Mark Stoney

TABLE OF CONTENTS

Accessing Security Incident Information on the Internet

M. A. Belsis[1], A. N. Godwin[2], L. Smalov[3]
Data Knowledge Engineering Research Group
Coventry University
Coventry, U.K.

[1]*E-mail: Belsis@coventry.ac.uk*
[2]*Email: a.m.godwin@coventry.ac.uk*
[3]*Email: l.smalov@coventry.ac.uk*

ABSTRACT

Computer security Incident Response teams have emerged due to the increase of computer crime. These can be national, international or organisation based. Maintaining a CSIRT poses a number of problems .In this paper the authors describe two of the technical problems that CSIRT's have, the storage and the acquisition of incident data. The paper describes a system based on the CORBA model that can be used for the efficient management of the incident recording database. The proposal also provides for alternative ways of accessing the database by companies and security analysts.

Keywords: Computer Security Incident Response Teams, Security Incident reporting and retrieval.

INTRODUCTION

The current rise of computer crime has sought the need for better information security (power 2000) and (Icove et al 1995). To assist organisations Computer Security Incident Response Teams (CSIRTs) have been assembled. The job of these teams is to receive information of computer security incidents that take place, analyse them and propose solutions to organisations. Some of the CSIRTs go further and assist companies in identifying the perpetrators and prosecuting them. CSIRTs use large databases that record details of security incidents coming from a wide variety and size organisations. They need to maintain information on this experience to develop defensive strategies, as described in (Anderson 1994).

Delivering a working CSIRT includes providing solutions to a number of problems. Some of these problems involve the management of the team. Examples include deciding on the boundary, the structure and the policy the team will follow. Currently numerous documents have been written on the way a CSIRT can be developed (NIST 1991), (Moira et al 1998) and (Brownlee 1998). These documents assist the teams in developing an overall policy and structure. Problems at a technical level include the acquisition, storage and analysis of security incident related information

This paper discusses the problems that are inherent in the acquisition and storage of security incident related information. These problems are briefly explained and short descriptions of current solutions are provided. Although there is much discussion on the design of common security incident data models there is little discussion on the design of systems to support the proposed model. This paper proposes the use of the CORBA model in the process of acquiring security incidents related information.

SECURITY INCIDENT REPORTING STRUCTURES

A number of CSIRT's exist; examples of such are the CERT/CC, CIAC and the CIRDB from CERIAS Laboratory. Each of them has developed and uses their own data model to organize the reported security incidents. Generally those data structures are focussed on storing the technical details of an attack. Modern trends in hacking involve numerous hosts and/or networks located around the world (e.g. distributed denial of service attack). To be able to trace information about attackers involved in such attacks CSIRT's need to collaborate (Athman et al 1999a and 1999b) . In order for the CSIRT's to collaborate they need to exchange precise incident information. The current incident recording structures in use this can be extremely difficult. CSIRT's store different types of information relating to security incidents. This means that two CSIRT's will often exchange information by telephone, fax or email.

Currently a lot of work is being done in developing common data models of security incident and/or attacks. An example of such is the European proposal, Project *S2003*, which proposes a simple incident data structure and gives guidelines on its use in developing a comprehensive library of security incidents (EU 1992). The proposed model can be used by European Computer Security Incident Response Teams (CSIRT) as the means of storing data collected from security incidents.

An interesting approach is the Incident Object Description and Exchange Format (IODEF) developed by the Incident Taxonomy and Description Working Group (TF-CSIRT) (Demchenko 2001) The model was created to assist CSIRT's to exchange incident data. The model is based on the IDEFDM (Herve 2000) model and so it can automatically read incident information from any intrusion detection sensor that uses the IDEFDM model through the use of XML. Although the model was first created as a mechanism for information exchange it can be also used to develop a CSIRT Database.

The authors of this paper have delivered their own proposal in this area (Meletis et al 2002). This structure is part of work in progress aiming to provide a General Enterprise IT Security Data Model.

Deciding on a common agreed structure is not a complete solution. The common structure will solve only some of the problems that CSIRT's have. Other issues are concerned with the way this structure will be used, in particular the way data are acquired, protected and retrieved.

ISSUES ON SECURITY INCIDENTS REPORTING AND RETRIEVAL

The security incident data model design should provide sufficient scope for the storage of data to satisfy requirements of all classes of its users. These users can be categorised crudely as 'security manager' and 'incident reporter' and 'incident analyst'. Each of these categories has its own needs in terms of scope and means of access. All will have a requirement in relation to the three standard three aspects of security infrastructure; namely: Confidentiality; Integrity, and Availability (Ravi et al 1991) and (Sushil 1996).

Confidentiality is vital for any incident database. Due to the nature of its data, such a database could be an invaluable tool for all kinds of hacker or criminal. Adversaries might be able to identify precisely the hardware and software that a company uses. The database might be used as an on-line hacking tutorial to advice adversaries on how to break into specific systems. The problem becomes more apparent if the database is going to be accessed via the Web. In order to trust the database, there must be security procedures in place to protect data stored in it.

Currently most CSIRTs use the World Wide Web as the medium to provide access to their databases. Due to the requirements of security in the WEB context, CSIRT's provide users with only a narrow view of the details that the incident includes. Usually this deals with the low level technical details of the attacks used and countermeasures for the attack. These details omit relationships to particular features within the target system. Although helpful to the technical expert they are far from useful to the manager of a company. Corporate managers need to identify the managerial information related to

an incident. Examples of such are: an average cost to the enterprise; the time the company needed to recover from the incident, and, statistical data. The statistics will relate to the frequency of the attack and/or the type of companies this incident targets. This information will assist managers in identifying potential 'need to secure' points and allow them to calculate a budget and/or extend the organisation's *high level* security policy. For the security manager it would not be helpful to display a huge amount of technical information on the screen when data incidents are retrieved. The need for security limits the utility of the current systems.

For current systems the retrieval of information takes place through the use of a simple search engine, using the name or code for the incident as the search key. The current data retrieval engines used by CSIRT's do not allow users to build their own *smart* queries. Security experts and security managers requiring access to the database should not need to become expert in the structure of the security incident database.

The above requirements of security managers raise issues in relation to the operation of a CSIRT database as well as issues for the scope of the data model for the database structure. In addition to having a data model that provides sufficient scope, the way the content is accessed must be appropriate to the user. Examples of smart queries that would be useful to both a security expert and a manager include:

- How many security incidents involved an internal user?
- How many security incidents targeted a buffer overflow in the Apache Web Server?

As noted above the answers for the different category of user might need to be different. The software behind the database should create views of the data targeting different classes of user. Example views could include: The *Management view,* and, the *Technical view.*

The above discussion has focussed on issues associated with retrieval. The high requirements of security, the incident reporting process posses force the CSIRT's to adopt manual or a semi-automated mechanism for capturing incident descriptions. To enhance security CSIRT's use techniques such as telephone, Fax or emails to gather information of security incidents. These can make the process of registering an incident slow and also make technically oriented personnel uncomfortable. An example of this can be seen in (CERT 2000) , where the form used by the CERT organisation is shown. It allows companies to register their computer security incidents, using FAX or email.

THE CORBA MODEL

The *Object Management Group* (OMG) introduced *Common Object Request Broker Architecture* (CORBA) in 1990. The CORBA model (fig.1 (OMG 1998c)) was aimed at providing an environment where software products, from different vendors could work together (OMG 1998c). In CORBA environment software products can exchange data and processes regardless their manufacturer, their operating system and the way they have been built.

Although CORBA aims at providing a standardised object oriented environment the programs that use this model do not have to be fully object oriented (i.e. Pseudo objects can be created for non-object software implementations).

CORBA does not provide ways of implementing a software package but introduces a way of creating an interface so other programs can access it. In order to build such an interface CORBA introduces an interface language called *Interface Definition Language* (IDL).

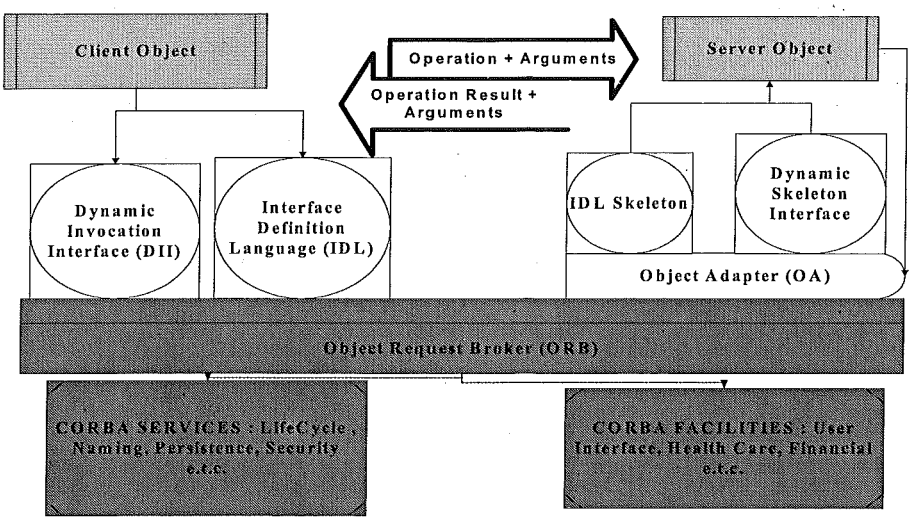


Figure 1: The CORBA Model

The CORBA/Java combination is the heart of the OMG's *Object Vision* (Orfali 1997). This combination allows the development of smart applications, that can locate the object services they need anywhere around the Internet without considering issues such as O/S, platform or implementation language. The applications of this type are based on component architecture. This allows software vendors to implement, modify or remove components from an application with a great deal of flexibility. Java applets can use the CORBA's *Dynamic Invocation Interface* to identify new server objects and generate server requests "on-the-fly". The session between the Java applet and the CORBA server objects will persist until either side decides to disconnect.

Current research (Philippe et al 1996) uses the common gateway interfaces (CGI) to allow clients to access and invoke CORBA objects located on the WEB Server machine. To enhance interoperability of the data structures the model uses, OMG have tried to incorporate XML into the model (Douglas 2001) . XML is used in the description of the server objects as well as to describe the information servers and clients are exchanging.

To fulfil the security requirements that a distributed system has CORBA provides a description of a Security Service. The Security Specification establishes the *Security Objects'* needed functionality along with the relations between them. The document in (OMG 1997) describes the functionality that a CORBA secure implementation must provide. By using this as a framework and following the guidelines included in the specification the CORBA vendors can develop security services that not only can provide adequate security but are going to be able to intercommunicate. Currently a number of known security protocols have been used to provide secure communications in a CORBA environment. Examples of these include Kerberos, Sesame and SSL.

PROPOSED SYSTEM.

Currently a number of research publications have proposed the use of the CORBA model to access web based databases (Athman et al 1999a and 1999b) and (Ebru et al 1995). The system briefly discussed here uses the CORBA model to allow registered organisations to fully access a CSIRT database. The new system allows access through the TCP/IP protocol. In addition to this the new system automates the process of recording an incident by providing the ability to implement the client of the system as part of the overall company's security management console.

Every organisation that wishes to acquire help with its security incidents will need to pre-register with the CSIRT. The first time a registered security administrator enters the CSIRT Web Site he/she will

authenticate himself/herself to the server with the use of an X.509 digital certificate. The user will then download a java applet. After setting up the applet, the applet will open a connection to the CSIRT database server. From there the user will be able to update his/her company's incident record and/or perform queries to the database (Fig. 2). To allow more security the digital certificate will contain the privileges that this user has over the database.

To provide friendly access a Natural Language Interface Database (NLIDB) system e.g. (Androutsopoulos et al 1995) and (Ott 1992) is used. These systems provide the ability to use regular English expressions to search the database instead of SQL queries. Such a system is responsible for translating the English expression into an appropriate SQL statement and for formatting the search output into an acceptable form. Initial designs for the MS SQL Server and MS English Query have been constructed and the examples given in section 3 of this paper have been shown to be relevant to the proposed architecture.

The NLIDB server can create different views depending on the type of the user (i.e. manager or technical personnel). This could add a second layer of security to the system due to the fact that we can program the server to hide the result from 'confidential' fields.

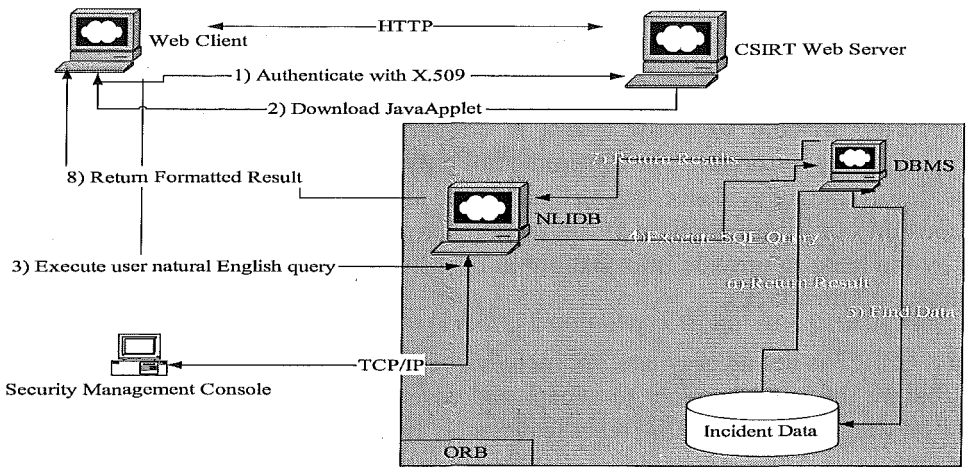


Figure 2: Proposed System

Another advantage of this model is that it allows interconnection with the existent security management consoles. Due to the open architecture that the CORBA model provides security vendors are able to incorporate functions that will allow their product to access the database without the use of the WWW. This will enable security managers and architectures to maintain distributed records of their systems security history. This provides security systems with the ability of automatically registering a security incident as soon as it happens (i.e. intrusion detection sensors can record an anomaly as soon as they detect one). In addition the CSIRT could update the intrusion detection sensors and firewalls of the registered companies to detect and stop the new anomaly. To be able to provide this functionality there are a number of problems that need to be solved first. Examples are: deciding on a common incident structure, and, ensuring that the detected incidents are not false.

The client software of the proposed system, either in the form of java applets or embedded code in the security management consoles, will be able to use CORBA's *DII* to identify and locate new services. CORBA will allow CSIRTs to add new functionality on demand.

An additional advantage of the proposed system is that although CSIRTs may use different data structures to store security incidents their DBMS systems will be able to interconnect and exchange information (Fig. 3). By interconnecting they will provide registered companies with a network of interconnected databases. This will provide a better variety of information that will enhance the

awareness of companies' employees and the more efficient production of security related statistical information.

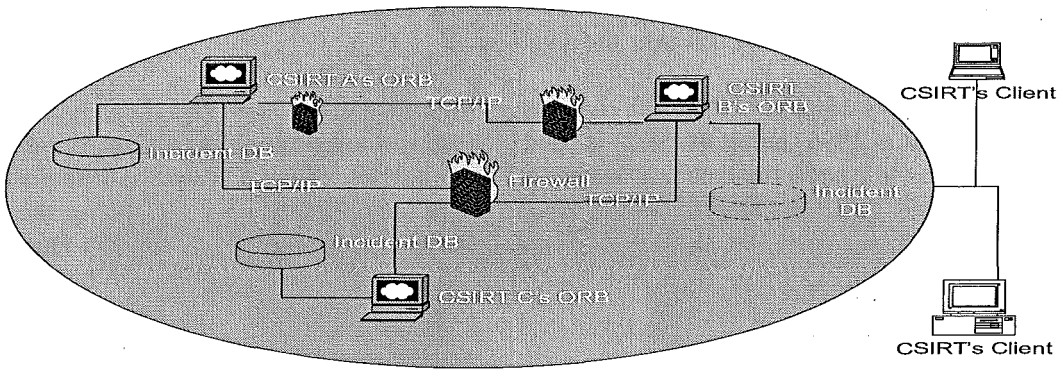


Figure 3: Interconnecting CSIRT

CONCLUSIONS.

IT attacks are currently on the increase. They are a significant drawback to the Internet's evolution. A Computer Security Incident Response Team is one of the best weapons that the IT community possesses against cyber terrorism. They promote not only security awareness to the IT community but assist companies in tracing evidence that will aid in the arrest of the attackers. In addition to that they provide a repository of information on attack techniques and their countermeasures. This information is analysed by the team to provide information on new trends in attacking as well as patterns that can be used to detect future attacks.

Developing a CSIRT is not an easy task. It includes both technical and managerial issues that have to be resolved before the team is ready to accept and assist in any security incidents. This paper has concentrated on describing two of the technical issues: Incident data structure, data Acquisition, and, data accessibility.

The system briefly described in this paper uses the CORBA model to allow an automatic registration and acquisition of security incident related information to take place. Additionally, the system allows the creation and use of smart queries by the CSIRT's personnel as well as by the client organisations. The use of the CORBA model provides the effectiveness of implementing the system along with the security management console that organisations use today and/or as part of the CSIRT web site. CORBA implementations provide a security service that is adequate in fulfilling the security requirements that the proposed system has. The incident database structure to be used must ensure that private and public information are stored in separated files having the private ones encrypted at all times.

CSIRTs that implement the proposed model will allow more automatic interoperability to take place between them. This will promote even more security awareness and collaboration.

REFERENCES

Anderson R., (1994), Why Cryptosystems fail, *Comm.* of the ACM, v. 37, n. p.32-40.

Androutsopoulos I., Ritchie G.D., and Thanisch P. (1995), Natural Language Interfaces to DBs - An Introduction. Natural Language Engineering, vol. 1, part 1. Cambridge University Press, pp.29-81,

Athman Bouguettaya, Boualem Benatallah, Lily Hendra, James Beard and Kevin Smith and Mourad Ouzzani. (1999), World Wide Database- Integrating the Web, CORBA and Databases. Proceedings of the SIGMOD Conference ,pp. 594-596.

Athman Bouguettaya, Boualem Benatallah , Mourad Ouzzani and Lily Hendra. (1999), Using Java and CORBA for Implementing Internet Databases. Proceedings of the 15th International Conference on Data Engineering ,pp.218-227.

CERT, (2000) Coordination Centre, Incident Reporting Form.

Commission of the European Communities Security Investigations Projects, (1992), Project S2003-Incident Reporting a European Structure "Final Feasibility and Strategy Report". Report No19733.Version 1.0.

Demchenko Y. (2001), Incident Object Description and Exchange Format Data Model and Extensible Markup Language (XML), Internet Draft.

Douglas C. Schmidt and Steve Vinoski, (2001), Object Interconnections: CORBA and XML, Part 1: Versioning, C/C++ Users Journal.

Ebru Killic, Gokhan Ozhan, Cevdet Dengi, Nihan Kesim, Pinal Koksal and Asuman Dogac, (1995), Experiences in using CORBA for a Multidatabase Implementation. In Proc. Of 6th Intl. Workshop on Database and Expert System Applications, London.

Herve Debar, Ming-Yuh Huang, David J. Donahoo, (2000) Intrusion Detection Exchange Format Data Model (IDEFDM).

Icove D., Seger K. and VonStorch W., (1995), Computer Crime: A Crimefighter's Handbook, O'Reilly Inc.

Meletis A. Belsis, Nick Godwin, Leon Smalov, (2002), A Security Incident Data Model, submitted to the 17th International Conference on Information Security, Egypt.

Moira J. West-Brown, Don Stikvoort, Klaus-Peter Kossakowski, (1998), Handbook for Computer Security Incident Response Teams (CSIRTs), CMU/SEI-98-HB-001.

N. Brownlee, (1998), Expectations for Computer Security Incident Response, RFC 2350.

NIST, (1991), Establishing a computer security incident response capability. Special publication 800-3.

Object Management Group, (1997), the Common Object Request Broker: Security Specification. 492 Old Connecticut Path, Framingham. MA 01701. USA.

Object Management Group, (1998c), The Common Object Request Broker: Architecture and Specification Revision 2.2. 492 Old Connecticut Path, Framingham. MA 01701. USA .

Orfali, R., Harkey, D. and Edwards, J. (1997), Instant CORBA, John Wiley and Sons Inc.

Ott N., (1992), Aspects of the automatic Generation of SQL Statements in a Natural Language Query Interface, Information Systems, 17(2),pp.147-159.

Philippe Blackbird, Christophe Gransart, Jean-Marc Geib, (1996), CorbaWeb: In WWW and CORBA Worlds Integration. Advanced Topics Workshop on Distributed Object Computing one the Internet, in

conjunction with 2nd Conference one Object-Oriented Technologies and Systems (COOTS), Toronto, Canada.

Power R,(2000), CSI/ FBI Computer Crime and Security Survey, Computer Security Issues and Trends, Vol. VI, No. 1.

Ravi S. Sandhu and Sushil Jajodia, (1991), Honest Database that can keep Secrets, Proc. Of the 14th NIST-NCSC National Computer Security Conference, Washington D.C., pp. 267-282.

Sushil Jajodia, (1996), Database Security and Privacy, ACM Computing Surveys, Vol. 28, No.1,

An Analysis of Public Key Cryptosystems

Y. M. Bani Hammad

Edith Cowan University, Australia
Email: ymubaraki@islamway.net

ABSTRACT

The research is based on analysis of public key cryptosystems. There are various public key systems that, in the past 25 years, have been designed with security, which depends on the difficulty of factoring large numbers. These include the RSA algorithm, Okamoto-Uchiyama, MultiPrime, etc. To date most factoring methods have been applied to numbers, which are the product of two primes. Recently, there have been several new public key algorithms proposed whose underlying modulus is the product of more than two primes. The aim of this research is to conduct an in depth analysis of methods of factoring such numbers.

Keywords: *cryptography, Computer Security, Public key, Factorization.*

INTRODUCTION

The first notions about security in computer systems came from military circles and were especially aimed at guarding the secrecy or confidentiality of data in information systems [JS95]. There have been, until recently, two main categories of cryptography: symmetric and asymmetric. Symmetric cryptography began about four thousand years ago. Asymmetric encryption was begun in 1976 by Diffie, and Hellman and it is known as the public key infrastructure (PKI). Asymmetric cryptography has led to a new trend of using cryptography and has solved some symmetric problems such key exchange and key management. In addition, it added some features to the application such as digital signature. Symmetric encryption, also called private key or secret key, uses the same key for sending and receiving [J98] as shown in Figure 1.

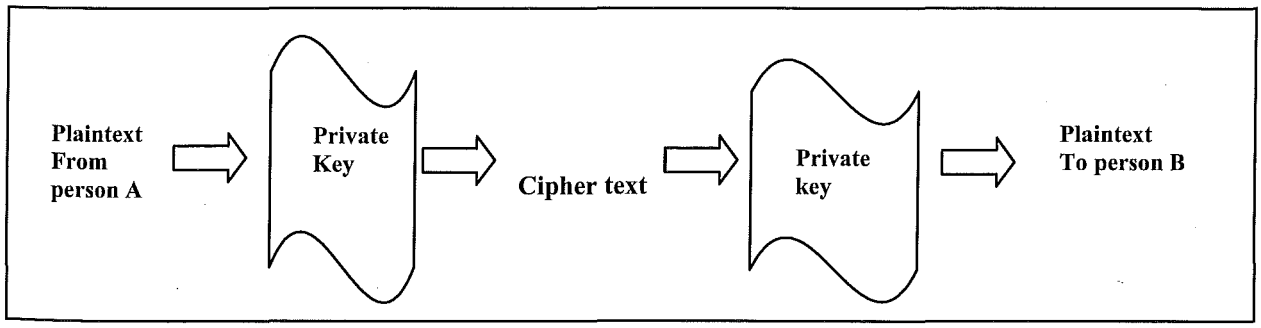


Figure 1.

Asymmetric encryption (PKI) uses two keys, one private and one public. In most cases, the public key is used for encryption and the private key is used for decryption. As shown in Figure 2

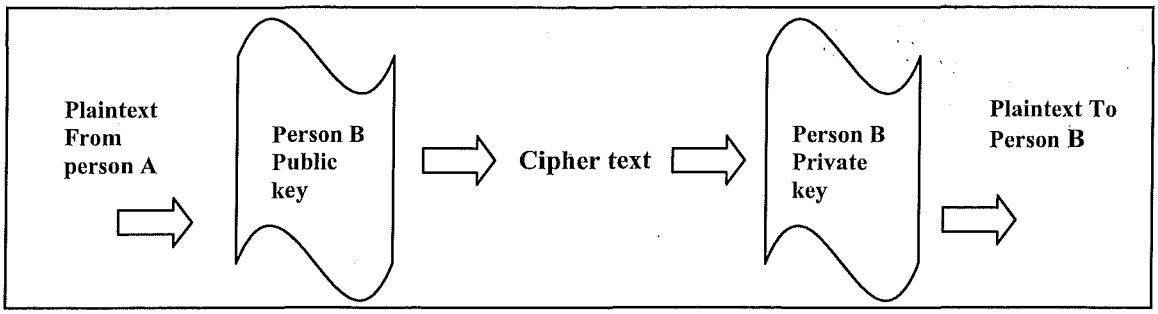


Figure 2: Public key for encryption and private key for decryption

Thus, the PKI has been the most important advance in cryptography [HD00] in four thousand years [CRYH]. Nowadays people use asymmetric encryption to exchange the key between two parties and then they encrypt the message within a symmetric encryption system. With this combination of symmetric and asymmetric cryptography, symmetric keys are typically ephemeral, that is, used once they are discarded. Symmetric keys are used for bulk encryption [AWCD01].

There are several algorithms for public key: One is RSA, named after its designers, Rivest, Shamir and Adelman. It was developed in 1978 [C97], and remains the most widespread public key algorithm in use today [AWCD01].

The Okamoto-Uchiyama public key cryptosystem uses $n = p^2q$ instead of $n = pq$ in the RSA public key. Therefore, the encryption speed of Okamoto-Uchiyama scheme is slower than that of RSA, however, the decryption speed is faster than that of RSA [OU98]. In MultiPrime public key cryptosystem $n = p_1 p_2 \ldots p_k$ where k is an integer number greater than two. In RSA if p and q are 512-bit numbers, then $n = pq$ is 1024 bits. In MultiPrime system, for example when $k = 4$ (i.e. $n = p_1 p_2 p_3 p_4$) p_1, p_2, p_3, and p_4 of 256 bits also result in a 1024-bit modulus.

The number of digits of the numbers we can factor is about three and half times as large as 24 years ago. A 45 decimal digits number was the biggest number factored in 1978, in year 2002 the biggest number factored contains 158 decimal digits. It is important to observe that the algorithm is the most important issue in factoring.

To date the largest RSA number has been factored is RSA-512 bits, which is 155 decimal digits, and the RSA-576, which is 174 decimal digits, is the smallest RSA Challenge number today:

188198812920607963838697239461650439807163563379417382700763356422988859715234665485319060606504743045317388011303396716199692321205734031879550656996221305168759307650257059

RSA, Okamoto-Uchiyama and MultiPrime security depends on the difficulty of the factoring problem. The difference between those algorithms in relationship of factoring is number of factors. RSA $n = pq$, Okamoto-Uchiyama $n = p^2q$, and Multiprime $n = p_1 p_2 p_3 \ldots p_k$, where k is an integer greater than 2

In Okamoto-Uchiyama method when $n = p^2q$, we can assume that $a = p^2$, so $n = aq$, n becomes the product of primes. So, if we can find a and q, it is easy to determine the value of p.

In the MultiPrime public key, to provide the same level of security of RSA, when p and q are, for example, 512-bit numbers, then $n = pq$ is 1024 bits. In the Multiprime system when $k = 4$ (i.e. $n = p_1 p_2 p_3 p_4$) p_1, p2, p3 and p_4 of 256 bits also result in 1024-bit modulus. Therefore, to find all factors of n first, we find one of them and let $n = p_1 p_2 p_3 p_4$, $a = p_1 p_2$, $b = p_3 p_4$, so $n = ab$, and we can implement in parallel system to find a and b.

The most straightforward method of factoring is trail division, where one simply tries to divide by each prime up to the square root of the number to factor. This method is indeed guaranteed to find a

factor of any composite integer, but it is also guaranteed to be computationally infeasible for large enough integers [Mat98]. PKI security is based on the mathematical property that factoring numbers that are products of two or more large primes (i.e. integers with several hundred decimal digits [DM87]) in a large finite field [A01], and solving discrete log problems, and this is difficult.

THE CURRENT SITUATION

The most popular asymmetric (or public key) is RSA. There have been many attempts to crack RSA. RSA Security Inc. has announced the results as follows [RSA]:

In April 1994, an international cooperative group of mathematicians and computer scientists solved a 17-year-old challenge problem, the factoring of a 129-digit number, called RSA-129, into two primes.

That is,
RSA-129 =
114381625757888867669235779976146612010218 2
96721242362562561842935706935245733897830 59
71235639587050589890751475992900268795435 41
=349052951084765094914784961990389813341776463849338784399082057 7times
32769132993266709549961988190834 46141317764296799294253979828853 3

The project took eight months and the equivalent of approximately 750 ten-MIPS computers [J98]. The team used over 6000 computers on the Internet running the quadratic sieve algorithm for eight months to factor the number.
In February 1999, a group of researchers completed the factorization of the 140-digit RSA challenges Number [RSA140].

In August 1999, a group of researchers completed the factorization of the 155-digit (512 bit) RSA Challenge Number. The work was accomplished with the General Number Field Sieve [RSA155].

The largest integer factored with a general algorithm is $2^{953}+1$, whose 158-digit cofactor was factored into a 73- digit factor and a 86-digit factor by Bahr, Frank and Kleinjung on January 18, 2002 [LRGS]

The following table illustrates the work done to factorize a number.

Year	Size	Number	Who	Method	Hardware
1970	39	$2^{128}+1$	Brillhart / Morrison	CFRAC	IBM Mainframe
1978	45	$2^{223}-1$	Wunderlich	CFRAC	IBM Mainframe
1981	47	$3^{225}-1$	Gerver	QS	HP-3000
1982	51	$5^{91}-1$	Wagstaff	CFRAC	IBM Mainframe
1983	63	$11^{93}+1$	Davis / Holdridge	QS	Cray
1984	71	$10^{71}-1$	Davis / Holdridge	QS	Cray
1986	87	$5^{128}+1$	Silverman	MPQS	LAN Sun-3's
1987	90	$5^{160}+1$	Silverman	MPQS	LAN Sun-3's
1988	100	$11^{104}+1$	Internet	MPQS	Distributed
1990	111	$2^{484}+1$	Lenstra / Manasse	MPQS	Distributed
1991	116	$10^{142}+1$	Lenstra / Manasse	MPQS	Distributed
1994	129	RSA-129	Atkins	MPQS	Distributed
1996	130	RSA-130	Montgomery	GNFS	Distributed
1998	140	RSA-140	Montgomery	GNFS	Distributed
1999	155	RSA-155	Montgomery	GNFS	Distributed
2002	158	$2^{953}+1$			

Table 1-Historical Factoring Record [HIS]

Current RSA Challenge numbers are the following [NUM]:

Challenge Number	Prize ($ US)	Status
RSA-576	10,000	Not Factored
RSA-640	20,000	Not Factored
RSA-704	30,000	Not Factored
RSA-768	50,000	Not factored
RSA-896	75,000	Not factored
RSA-1024	100,000	Not factored
RSA-1536	150,000	Not factored
RSA-2048	200,000	Not factored

Table 2- RSA Challenge Number

There is a relationship between the private and public key, everyone is able to know what the public key for another person is. So it follows that some organizations may know how to determine a private key from a public key.

Therefore, public and private keys must be used carefully. People must use separate methods and/or keys for encryption and signing, and they must never use their private key to sign the exact message sent to them and they should never sign unknown messages.

The most obvious way to break an algorithm is to try every possible key. This is a well-known, brute force, exhaustive search method. It requires a very small chunk of ciphertext and an equally small chunk of corresponding plaintext. This attack is always possible and there is no way to prevent it. The best one can do is to make the attack too expensive, both in time and money [B95]. For example if we want to crack a 40-bit key, the total number of possible keys will be $2^{40} = 10^{12}$.

No one has proven mathematically that the security of RSA is dependent on the difficulty in factoring large numbers. It is only conjectured. Everyone believes that the only way to break RSA is to factor the large number, but it is always possible that someone could discover another way [B95]. To date there is not known way to break the RSA system without factoring n [G92]. There are several methods that a hacker can use to crack a code, including: Known plaintext attack which is an attack based on given plaintext and the corresponding ciphertext [JV98]; Man-in-the-middle, where the hacker is hidden between two parties and impersonates each of them [W99]; Active attack, where the hacker inserts or modifies messages; Cut and paste, where the hacker mixes parts of two different encrypted messages and, sometimes, is able to create a new message. This message is likely to make no sense, but may trick the receiver into doing something that helps the hacker [W99]; Time resetting; some encryption schemes use the time of the computer to create the key. Resetting this time or determining the time that the message will be created can give some useful information to the hacker [W99]. Some algorithms are only breakable with the benefit of more time than the universe has been in existence for and a computer larger than all the matter in the universe. Such algorithms are theoretically breakable, but not breakable in practice. An algorithm that is not breakable in practice is secure [B93]. On the other hand, if someone cannot break some algorithm, it does not mean that this algorithm cannot be cracked.

FACTORIZATION METHODS

The number of digits of the numbers we can factorise is about three and half times as large as 24 years ago. Since RSA publication in 1987, efforts to break the cipher have resulted in increased activity in developing faster and better factorisation algorithms for integers [JL90]. It is important to observe that the algorithm is the most important issue in factoring. Development of factorisation algorithms is a great interest and research on this is being done all over the world. The algorithms are getting better,

but no one has found a very good algorithm. On the other hand, no one has succeeded in proving that a good algorithm does not exist [NADA]. Below are some algorithms used to factor a number:

Fermat Factorisation

There is a way to factor a composite number n that is efficient if n is a product of two integers, which are close to one another, this method is based on the fact that n is then equal to a difference of two squares, one of which is very small [N91]. To formalize this as an algorithm, we take trial values of the number a from the sequence int(sqrt(n)), int(sqrt(n)+1, ... and check whether a^2 - -n is a square. If it is, say b^2, then $n = a^2 - b^2 = (a+b)(a-b)$. Many of the successful factoring methods of the past twenty years have used this same basic technique, which dates back to the time of Fermat. Appendix B illustrates a numeric example of factoring a composite number using the Fermat method.

Appendix A illustrates an innovative technique to factor a product of two primes based on equation (7). This method is faster than the Fermat method,.

Pollard Rho Number Factoring

Also called the "Monte Carlo" method [N91]. This algorithm was developed by John M. Pollard in 1975[RCP00], and improved upon by R. P. Brent [POL]

It uses the following algorithm [RHO]

Choose a simple polynomial with integer coefficients, such as $f(x) = x^2 + 1$. Next, choose some particular value $x = x_0$ (perhaps $x_0 = 1$ or 2, or perhaps a randomly generated integer) and compute the successive iterates of f: $x_1 = f(x_0)$, $x_2 = f(f(x_0))$, $x_3 = f(f(f(x_0)))$, etc. That is define $x_{j+1} = f(x_j)$, j = 0,1,2,3...

Then make a comparison between the different x_j's, to find two that are in different residue classes modulo n, but in the same residue class modulo some divisor of n. Once we find such an x_j, x_k we have gcd ($x_j - x_k$, n) equal to a proper divisor of n, and we have completed the task [N91]. This method depends of two things, initial random value (x_0), and function f.). Appendix B illustrates a numeric example of factoring a composite number using Pollard Rho method.

Factor Base Factoring Method

A Factor Base is a set $B = \{-1, p_1, \ldots, p_h\}$, where $\{p_1, \ldots, p_h\}$ is a set of prime numbers. We say that the square of an integer b is a B-number (for a given n) if the smallest absolute residue b^2 mod n can be written as a product of numbers from B [N91].Factor Base is a generalization of the idea behind Fermat factorisation, leading to a much more efficient method. It states that when n is a composite number, and we can determine some t and s of the form $t^2 \equiv s^2 \bmod n$, when $t \neq \pm s \bmod n$, we immediately find a factor of n by computing gcd (t+s, n) or gcd (t-s, n).

Continued Fraction Factoring Method

Originally proposed by Legendre [N91], finds many b such that $b^2 \bmod n < 2\text{sqrt}(n)$. The following illustrates the Continued Fraction algorithm:

Let n be the integer to be factored.
First set $b_{-1} = 1$, $b_0 = a_0 = [\text{sqrt}(n)]$, and $x_0 = \text{sqrt}(n) - a_0$.
Compute $b_0^2 \bmod n$. Next, for i = 1,2... do the following:

1- Set $a_i = [1/x_{i-1}]$ and then $x_i = 1/x_{i-1} - a_i$
2- Set $b_i = a_i b_{i-1} + b_{i-2} \bmod n$
3- Compute $b_i^2 \bmod n$. After doing this for several i, look at the numbers in step 3, which factor into ± product of small primes. Take factor base B including –1. Then list all of the numbers $b_i^2 \bmod n$, which are B-numbers, along with corresponding vectors V_i of zeros and ones, then find a subset whose vectors sum to zero. Set $\mathbf{b} = \sum b_i$. Set $\mathbf{c} = \sum p_j^{f_j}$, where p_j are the elements of B (except for –1) and $f_j = (1/2) \sum y_{ij}$. If $\mathbf{b} \neq \pm \mathbf{c} \bmod n$, then look for another subset of i such that $\sum V_i = 0$, then continue computing more a_i, b_i and $b_i^2 \bmod n$, enlarging factor base B if necessary.

The Elliptic Curve Method

[E99]: An elliptic curve is a set of point (x, y) satisfying $y^2 = x^3 + ax + b$, when $4a^3 + 27b^2 \neq 0$. We denote the point at infinity by O. There are five rules for adding points:

1- $O + O = O$
2- $(x,y) + O = (x,y)$
3- $(x,y) + (x,-y) = O$
4- If $x_1 \neq x_2$, $(x_1,y_1) + (x_2,y_2) = (x_3,y_3)$ where
   $x_3 = (r^2 - x_1 - x_2) \bmod p$
   $y_3 = (r(x_1 - x_3) - y_1) \bmod p$
   and $r = (y_2 - y_1) / (x_2 - x_1) \bmod p$
5- If $y \neq 0$, $(x,y) + (x,y) = 2(x,y) = (x_2, y_2)$
   Where
   $x_2 = (r^2 - 2x) \bmod p$
   $y_2 = (r(x - x_2) - y) \bmod p$
   and $r = (3x^2 + a) / (2y) \bmod p$

A key reason for the increasing interest in elliptic curves on the part of cryptographers is the ingenious use of elliptic curves by H.W. Lenstra to obtain a new factorisation method. Let n be a composite odd integer that we want to factorize.

First we generate pairs (E, P) consisting of an Elliptic curve $y^2 = x^3 + ax + b$ with $a, b \in Z$, and a point $P = (x, y) \in E$. Given such a pair, we go through the following procedures:

If that procedure fails to yield a nontrivial factor of n then we generate a new pair (E, P) and repeat the process.

Then, we calculate $4a^3 + 27 b^2$. If gcd $(4a^3 + 27 b^2, n) \neq 1$, we have a divisor of n, and the task is done. If gcd equals n then we must choose a different elliptic curve. If gcd equals one, we may proceed.

Next, we choose two positive integer bounds B and C. B is a bound for the prime divisors of the integer k by which we multiply the point P. if B is large, then there is a greater probability that our pair (E, P) has the property that $kP = O \bmod p$ for some p|n; on the other hand, the larger B the longer it will take to compute kP mod p. C is a bound for the prime divisors p|n for which we are at all likely to obtain a relation $kP \bmod p = O$

The Quadratic Sieve (QS)

The QS method for factoring large integer was developed by Pomerance in the 1981 [N91]. For a long time it was more successful than any other method in factoring integers n of general type which have no prime factor of order of magnitude significantly less than sqrt (n), and it is still the method of choice for integers between 50 to 100 digits [MAT98].

The General Number Field Sieve (GNFS)

GNFS was invented by John M. Pollard in 1988 and further developed in 1994 by J. P. Buhler, H. W. Lenstra Jr. and Carl Pomerance [NFS]. GNFS is the fastest [HIS] and the best-known algorithm for factoring large numbers is the General Number Field Sieve (GNFS). GNFS consists of a sieving phase that searches a fixed set of prime numbers for candidates that have a particular algebraic relationship, modulo the number to be factored. This is followed by a matrix-solving phase that creates a large matrix from the candidate values, and then solves it to determine the factors.

The sieving phase may be done in distributed fashion on a large number of processors simultaneously. The matrix-solving phase requires massive amounts of storage and is typically performed on a large supercomputer [BES]. It was used to factor the RSA-130, RSA-140, and RSA-155.

There are various public key systems designed in the past 25 years whose security depends on the difficulty of factoring large numbers. These include the RSA algorithm, Okamoto-Uchiyama, MultiPrime, etc. cryptosystem.

Fundamental theorem of arithmetic

Every positive integer n>1 can be expressed as a product of primes; this representation is unique, apart from the order in which the factors occur [DB89]. So, any positive integer n>1 can be written uniquely in a canonical form

n = p1k1 p2k2 ... prkr for i = 1,2,..., r , each k, is a positive integer and each pi is a prime with p1 < p2 < ...< pr

The problem now is how can we determine the prime factors of a given n. The following theorem, which we call RAK,* attempts to find the prime factors.

RAK theorem: let n be a product of two prime numbers p & q, n = pq. It is always true that xn+1 mod n = xp+q mod n, where gcd (x, n) =1

Proof:

n = pq
F (n) = n +1 - (p + q)
n + 1 = F (n) + p + q

From the Euler theorem which states that if gcd (x, n) =1, then $x^{\Phi(n)} \bmod n \equiv 1$[EULTH][WL76]. $x^{\Phi(n)} \bmod n \equiv 1$,

$$
\begin{aligned}
x^{n+1} \bmod n &\equiv x^{\Phi(n)+p+q} \bmod n \\
&\equiv x^{\Phi(n)} x^{p+q} \bmod n \\
&\equiv (x^{\Phi(n)} \bmod n)\,(x^{p+q} \bmod n), \text{ since } x^{\Phi(n)} \bmod n \equiv 1 \\
&\equiv x^{p+q} \bmod n
\end{aligned}
$$

Note: It is true that for any integer $x > 1$, $x^{n+1} \bmod n \equiv x^{p+q} \bmod n$. We select $x = 2$ to keep the calculation as small as we can.
It is easy to calculate $2^{n+1} \bmod n$ because n is a part of the public key. In addition, it always true that $(p+q) \geq \text{int}(\text{sqrt}(n) + 1) * 2$. So, to break an RSA number we have to determine p+q when we know $2^{p+q} \bmod n$. The hard part of this method is to determine the difference between $(p + q)$ and $\text{int}(\text{sqrt}(n) + 1) * 2$.

To factor a given n to p and q we take the following steps, which are based on the RAK theorem:

1- $a \equiv 2^{n+1} \bmod n$, so we know now that $2^{p+q} \bmod n \equiv a$
2- $b = (\text{int}(\text{sqrt}(n) + 1) * 2$, so $p+q = b + x$, in the following steps we look for x when x is an integer number, as x becomes, smaller we need fewer iteration to find p and q.
3- $e = \text{int}(\ln n / \ln 2)$, e will be our counter in the loop. So, we need to know the biggest value for e when $2^e < n$, to reduce the iterations number
4- $c \equiv 2^b \bmod n$
5- $d \equiv a * c^{-1} \bmod n$
6- If $\ln d / \ln 2 = $ integer number then $p + q = b + (\ln d / \ln 2)$ and the procedure is finished. If $\ln d / \ln 2 \neq$ integer number. Because we are looking for x when $2^x \bmod n \equiv d$, and when $x \leq \text{int}(\ln n / \ln 2)$ then $\ln d / \ln 2 = \text{int}$ go to step 7
7- $b = b + e$
8- $d \equiv d * (2^e)^{-1} \bmod n$, and go to step 6

Let r = the number of iterations that we need to factorise n, $e = \ln n / \ln 2$, and c is integer number $< e$. So,

$p + q = (\text{int}(\text{sqrt}(n) + 1) * 2) + re + c$
$x = re + c$
In Fermat method x will be $(p + q) - (\text{int}(\text{sqrt}(n) + 1) * 2)$

We can improve the speed of the previous technique by writing the following matrix x

The columns are (1,2,3,..., g) when $g = \text{int}(\ln n / \ln 2)$, and rows are (kg) when $k = (0,1,2,3,\ldots, g-1)$ in this case the speed of previous technique increased g times. We may note that the rows are not limited as we add more rows we increase the speed in each row by g.

	1	2	3	...	g
0	2	4	8	...	2^g
g	$2^{g+1} \bmod n$	$2^{g+2} \bmod n$	$2^{g+3} \bmod n$	...	$2^{2g} \bmod n$
2g	$2^{2g+1} \bmod n$	$2^{2g+2} \bmod n$	$2^{2g+3} \bmod n$	...	$2^{3g} \bmod n$
3g	$2^{3g+1} \bmod n$	$2^{3g+2} \bmod n$	$2^{3g+3} \bmod n$	...	$2^{4g} \bmod n$
...	...	...	...	...	...
g2 – g	$2^{(g\text{\^{}}2)-g+1} \bmod n$	$2^{(g\text{\^{}}2)-g+2} \bmod n$	$2^{(g\text{\^{}}2)-g+3} \bmod n$	...	$2^{(g\text{\^{}}2)} \bmod n$

$h = 2^{(g\text{\^{}}2)} \bmod n$
$j = h^{-1} \bmod n$
$k = jf \bmod n$ (*)
if k = any elements in matrix x let say x_{ij}, then $y = b + i + j$, y is p+q
if $k \neq$ any elements in matrix x then $f = k$, $b = b + g^2$ and go to * step

References

[A01] Aviel D.Rubin (2001), *White-Hat Security Arsenal,* Addison-Wesley

[AWCD01] Andrew Nash, William Duane, Celia Joseph, and Derek Brink (2001), *PKI: Implementing and Managing E-Security*, Osborne/McGraw-Hill

[B93] Bruce Schneier (1993), *Applied Cryptography*, John Wiley & Sons, Inc.,

[B95] Bruce Schneier (1995), *E-mail Security: How to Keep Your Electronic Message Private*, John Wiley & Sons, Inc.

[BES] http://www.rsasecurity.com/rsalabs/challenges/factoring/faq.html#WhatAreThe Best

[C97] Chales P.Pfleeger (1997), *Security in Computing,* Prentic Hall PTR

[CRYH] http://www.all.net/books/ip/Chap2-1.html

[DB89] David M.Burton (1989), *Elementary Number Theory,* Wm. C.Brown,

[DM87] Dennis Longley, and Michael Shain (1987), *Data & Computer Security Dictionary of standards concepts and terms,* Stockton Press

[EULTH] http://primes.utm/glossary/page.php/EulersTheorem.html

[G92] Gustavus J. Simmons (1992), *Contemporary Cryptography the Science of Information Integrity*, IEEE Press

[HD00] H.X Mel, Doris Baker (2000), *Cryptography Decrypted*, Addison-Wesley

[HIS] http://www.rsasecurity.com/rsalabs/bulletins/bulletin13.html

[J98] Jonathan Knudsen (1998), *Java Cryptography*, O'Reilly

[JL90] J.H. Loxton (1990), *Number Theory and Cryptography,* Cambridge University Press

[JS95] Jan H.P. Hllof and Sebatiaan H.Vou Solms (1995), *Information Security the Next Decade*, Chapman & Hall

[JV98] Jan C A, Van Der Lubbe (1998), *Basic Methods of Cryptography,* Cambridge University Press

[K96] Kevin S.McCurley (1996), *Advances in Cryptography AsiaCrypt 96*, Spriger

[LRGS] http://www.loria.fr/~zimmerma/records/factor.html

[N91] Neal Koblitz (1991), *A Course in Number Theory and Cryptography,* Springer

[NADA] http://www.nada.kth.se/~joel/qs.pdf

[NFS] http://www.frenchfries.net/paul/factoring/theory/nfs.html

[NUM] http://www.rsasecurity.com/rsalabs/challenges/factoring/numbers.html

[OU98] EUROCRYPT 98 (1998), *Advances in Cryptology EUROCRYPT '98*, Springer

[POL] http://www.frenchfries.net/paul/factoring/theory/pollard.rho.html

[RCP00] Richard Crandall and Carl Pomerance (2001), *Prime Numbers A Computational Perspective,* Springer

[RHO] http://mathforum.org/dr.math/problems/mcgrew10.26.98.html

[RSA] http://www.rsasecurity.om

[RSA40] http://www.ecst.csuchico.edu/~atman/Crypto/misc/rsa40-crack.html

[RSA48] http://news.com.com/2100-1023-269824.html?legacy=cnet

[RSA56] http://news.com.com/2100-1001-208505.html?legacy=cnet

[RSA140] http://www.rsasecurity.com/rsalabs/challenges/factoring/rsa140.html

[RSA155] http://www.rsasecurity.com/rsalabs/challenges/factoring/rsa155.html

[RSAFQ] http://www.rsasecurity.com/solutions/developers/total-solution/faq.html#IIIA

[W99] William Buchnan (1999), *Mastering Networks,* Macmillan Press LTD

[WL76] William W. Adams, Larry Joel Goldstein (1976), *Introduction to Number Theory,* Prentice-Hall, Inc.

A Method for Understanding Students' Perceptions of Concepts in the Defence in Depth Strategy

Clifton L Smith

Security and Applied Technology Research (SATR) Group
School of Engineering and Mathematics
Edith Cowan University, Western Australia
Email: clifton.smith@ecu.edu.au

Machine Vision Research Group
Department of Electrical and Electronic Engineering
Nottingham Trent University, England

ABSTRACT

The Defence in Depth strategy is a fundamental principle in the physical protection of the assets of an organisation. The robustness of the strategy has allowed it to be applied to a range of situations where assets need to be protected. This study seeks to examine the understanding of novice learners' perceptions of the defence in depth principles through the knowledge structure approach to concepts in the strategy. The multidimensional scaling (MDS) statistical technique has been applied to dissimilarity measures on a range of facilities according to the defence in depth functions of deterrence, detection, delay, and response. The barriers that correspond to these functions are considered as the analysis standard for the set of facilities.

Pre and post instruction knowledge structures have been developed for novice learners, and the knowledge structure for an expert group produced. The study indicates that novice learners knowledge structures become more like the experts structure with instruction.

Key Words: cognitive structures, knowledge structures, defence in depth, multi-dimensional scaling, MDS, understanding security concepts

INTRODUCTION

The *defence in depth* approach to the protection of assets is most familiar to security practitioners and operatives. The strategy has an extensive history of application and success in the prevention of theft, destruction of facilities, and the protection of personnel and information. Probabilistic models of the *defence in depth* principle have been developed to optimise the protection of assets in an organisation, and as a consequence the application of a range of barrier types to prevent unauthorised access is well understood.

Because of the robustness of the *defence in depth* principle, it has been applied in a variety of contexts ranging from physical security through to the protection of information. There is a need for students to understand the principle of *defence in depth* in its many applications, and be able to conduct appropriate analyses of penetration in the strategy. This study has conducted an investigation of a novel approach to the level of understanding of novices' perceptions of the concepts involved in the *defence in depth* principle.

The portrayal of knowledge structures for a discipline is an approach to research the *instruction – learning* paradigm for novice learners. Organisational and structural relationships among concepts may provide an understanding of the acquisition of concepts in memory, with a subsequent improvement in the instructional process. Although there are several approaches to the portrayal of knowledge structures, this study shall be concerned with those based on similarity, proximity or relatedness between concepts within the knowledge domain.

The paper describes a novel approach to understanding students' perceptions of concepts in a *defence in depth* strategy, and the development of these structures with instruction. Novice learner groups, both before and after instruction together with accumulated maturation with the *defence in depth* strategy, are compared to the knowledge structure of an expert group. The more alike are the knowledge structures of novice learners and the experts, then the more likely it is that learning has occurred.

DEFENCE IN DEPTH

The *defence in depth* strategy has been applied in the protection of assets for centuries, and the principle is based on the enclosure of an asset by a succession of barriers, all of which must be penetrated for the asset to be acquired. This approach to asset protection through a succession of barriers has been adopted as a strategy to restrict the penetration by unauthorised access to the assets, and hence gain time for the authorities to react to the penetration of the asset protection system (Smith and Robinson, 1999).

The functions of the barriers in the *defence in depth* strategy are those of deterrence, detection, delay, and response. These functions provide a range of types of barriers that maximise the probability of prevention of unauthorised access, and maximise the potential for detection and apprehension of unauthorised persons.

Deterrence: is achieved through signage, lighting, definitions of boundaries, and psychological cues.
Detection: if access has been gained by unauthorised person(s) then early detection of their presence is required to facilitate apprehension.

Delay: when unauthorised access to a facility has occurred, the physical barriers must be sufficiently substantial to delay the progress of the intruders.

Response: the delay time of the barriers to resist must be sufficient to allow an appropriate response team to attend the scene for apprehension.

These functions of the *defence in depth* strategy can be mapped onto the following types of barriers for the protection of assets (Smith and Robinson, 1999):

Psychological barriers: barriers that give clear warning to persons that a boundary should not be traversed, such as signs, lines, chains, fences, and lighting.

Electronic barriers: barriers to detect the presence of an intruder, such as CCTV, intrusion detection systems, and access control systems.

Physical barriers: the purpose of these barriers is to prevent physical access of the intruders to the asset, and include fences and walls, the envelope of the building, doors and windows, and safes and containers.
Procedural barriers: barriers that derive from security policy to maintain the integrity of the protection of assets, and include mobile guards, identification badges, sign in register, and proximity access control cards.

The defence in depth strategy has been operationalised to protect assets in facilities for a considerable duration, but has not been formalised with barrier types and functions. This study has investigated the perceptions of students of the major concepts of the *defence in depth* strategy in order to gain an insight into the understanding of this principle.

KNOWLEDGE STRUCTURES

The early cognitive researchers (Bruner, Goodnow, and Austin, 1956; Ausubel, 1966) have proposed that teaching effectiveness can be enhanced if the learner has memory structures appropriate for the instructional material. Cognitive models of human memory, such as information processing models, knowledge-assembly theory, and organisation theory all emphasise the necessity for structure in memory. Human long-term memory (LTM), like any large-scale data storage device, can have accessing and retrieval of information difficulties. A central aspect of the role of organisation of concepts in memory is considered to be related to the retrieval of information in recall.

The network of relations among concepts in LTM constitutes the substantive or subject matter structure of a particular cognitive domain, which is unique to the individual. The learning of the subject matter structure has been initially characterised by Michon (1972) as the acquisition of internal representations of external structures, and may be described by a network of relations between the concepts. A variety of studies using the structure of knowledge in a cognitive domain have been reported, including the fields of management, marketing, educational psychology, statistics, medicine, and physics (Gonzalvo, et al., 1994; Streveler, 1994; McGaghie, et al., 1998).

The task of the instructor is to assist with student acquisition into memory of the major concepts, so that the learner perceives *correct* relations between concepts of a discipline. It is assumed that the instructor knows the ideal organisation of concepts to be learned by a student, so that the function of *expert* instruction is to aid the learner to acquire the perceived knowledge structure of the instructor.

It can be argued that the experts of a discipline define the structural acquisition of concepts from the discipline; so that teachers, authors and researchers operating within a content domain perceive relations between concepts in a similar manner (Smith, 1986; Sireci and Geisinger, 1995). As the overall organisation of concepts in memory indicates the knowledge structure for a particular individual then a comparison of knowledge structures of learners and experts can reveal the extent of meaningful learning that has occurred. Koubek and Mountjoy (1991) showed this effect with subjects operating in the domain of clerical work, and Steinberg (1990) showed the novice – expert differences in knowledge structures in the domain of statistics.

MULTIDEMENSIONAL REPRESENTATION

Concepts can be *classified* by component properties, where the components of classification can be characterised by qualitative attributes called *features*, or quantitative attributes called *dimensions*. Thus a concept may be described by an extensive featural list indicating the attributes characteristic of the concept, or by a shorter list of dimensions which indicates how much of the attributes are present. Both approaches seek to indicate the degree of similarity between concepts (Smith, 1986).

This investigation uses the dimensional approach, which is a probabilistic representation of concepts, and where the dimensions are continuous attributes of the concept. Each dimension is represented as a continuous psychological dimension, with the different between two concepts being a matter of continuous degree on each defining dimension.

Concepts having the same relevant dimensions can be represented as points in a multidimensional metric space, where the defining dimensions are orthogonal. The relations between pairs of concepts

that occupy positions in multidimensional space can be expressed as a distance parameter. This parameter can be used as a measure of similarity between concepts with respect to the defining dimensions of the space.

METHODOLOGY

Multidimensional scaling (MDS) configurations are representations of subjects' perceptions of similarity between the concepts in the knowledge domain. This study selected a range of facility concepts to be judged according to types of barriers in the *defence in depth* principle. These types of barriers of psychological, electronic, physical, and procedural barriers are derived from the functions of deter, detect, delay, and respond. The facility concepts selected are*: military base, supermarket, art gallery, suburban bank, research centre, university, primary school.* These facility concepts were chosen as they represent a large range of dissimilarity on the dimensions of knowledge space for the domain of the *defence in depth* principle.

Students enrolled in the Security Science course at Edith Cowan University were administered a dissimilarity instrument that required that they judge the similarity/dissimilarity between pairs of facility concepts according to each of the types of barriers in the *defence in depth* principle. That is, these novice learners were required to indicate the degree of similarity for the physical security between the facilities, where Figure 1 shows Art Gallery – Primary School, and University – Military Base facility concept pairs for similarity/dissimilarity rating.

Physical Barriers

Unrelated Highly Related

Art Gallery – Primary School [----------]

University – Military Base [----------]

Figure 1: Similarity/Dissimilarity rating scale for spatial distribution of facility concepts.

Students indicated the degree of similarity between each of the twenty-one pairs of facilities generated for the seven selected facility concepts. These indications of similarity, or dissimilarity, for each of the pairs of facility were completed for the four types of barriers in the *defence in depth* strategy.

The *defence in depth* concepts similarity instrument was completed by the following groups:

ten novice Security Science degree level students before instruction in Defence in Depth.
ten novice Security Science degree level students after instruction in Defence in Depth.
four Master of Science (Security Science) candidates as experts in Defence in Depth.

The data from the subject groups were collated in both individual and group averaged data in the form of dissimilarity half-matrices of the pairs of facilities. These data were analysed to examine the knowledge structures as indicators of understanding in the domain of the *defence in depth* principle.

ANALYSES AND INTERPRETATION

The analytical methodology was the Multidimensional Scaling (MDS) technique (Takane, Young, and De Leeuw, 1977) to obtain the single most satisfactory representation of the distribution of the concept facilities according to the barrier types for the knowledge structures. These structures were derived in 2-D representations for the pre and post novice instruction groups and the expert group. From these research groups, both *within* and *between* comparisons were conducted for groups and individuals.

The following analyses were performed to test the hypothesis of differences of understanding between expert and novice knowledge structures in the *defence in depth* principle of asset protection. These analyses were individually conducted for all barrier types for:

ten novice learners pre instruction defence in depth
ten novice learners post instruction defence in depth
four post graduate students in Security Science

All pair-wise comparisons between concepts were coded as dissimilarities, and these data became input for multidimensional scaling analysis as representations of cognitive structure. Although all data were analysed for three dimensions and two dimensions solutions, only two dimensional structures have been presented. However, a limited selection of structures have been presented to enable discussion of the proposed technique.

The Figure 2 shows the 2-D knowledge structure for the novice pre instruction group for Physical Barriers, with the distribution of the facility concepts according to the dissimilarity ratings for this group. The spatial distribution of the concepts represents the knowledge structure, with the dimensions for the configurations undefined at this stage.

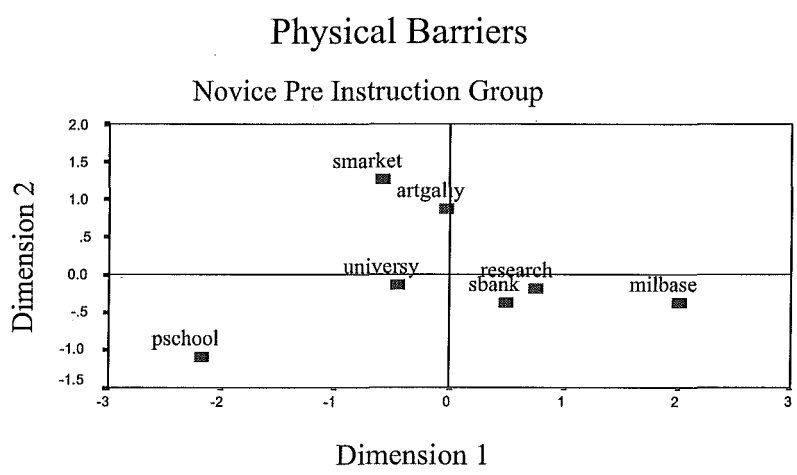


Figure 2: Knowledge structure for Physical Barriers for Novice Pre Instruction Group.

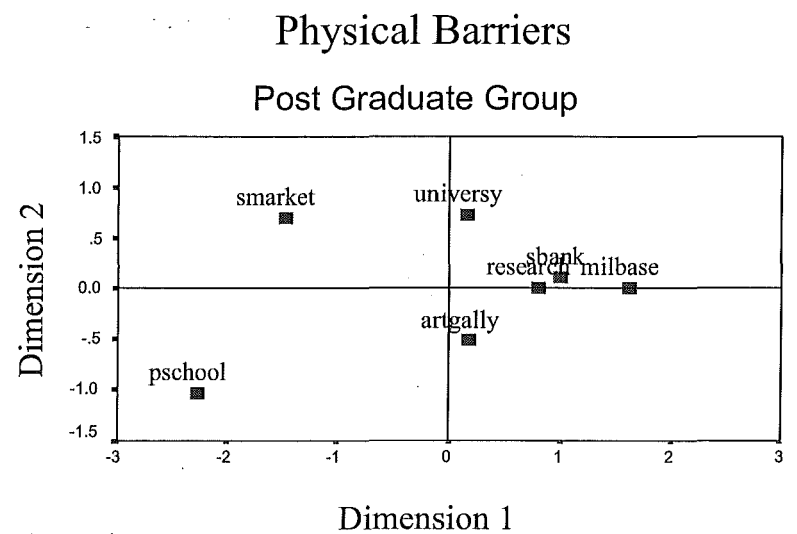


Figure 3: Knowledge structure for Physical Barriers for Post Graduate Group

However, the experts' group knowledge structure represents the formal structure of the content domain (Smith, 1986), and is the structure that novice learners aspire to achieve through understanding of the principle of *defence in depth*. The Figure 3 shows the expert group 2-D knowledge structure for Physical Barriers for the facility concepts. It is expected that this structure approaches that of the ideal configuration.

Although all four types of barriers for the selected facility concepts have been analysed, only knowledge structures for Procedural Barriers for the novice groups, pre and post instruction structures have been compared to the experts structure. The Figure 4 displays the knowledge structure of Procedural Barriers for the novice group before the instructional phase of the project. The relationships of the spatial distribution for the concepts have no reliability as diminished understanding of the concepts are evident as the novices had not studied the knowledge domain.

Following the instructional phase of the project when the learners were presented with the concepts and principles of Defence in Depth, the novice group again responded to the instrument producing dissimilarity data for spatial analysis. The Figure 5 shows the 2-D configuration for the facility concepts for Procedural Barriers in the Defence in Depth strategy.

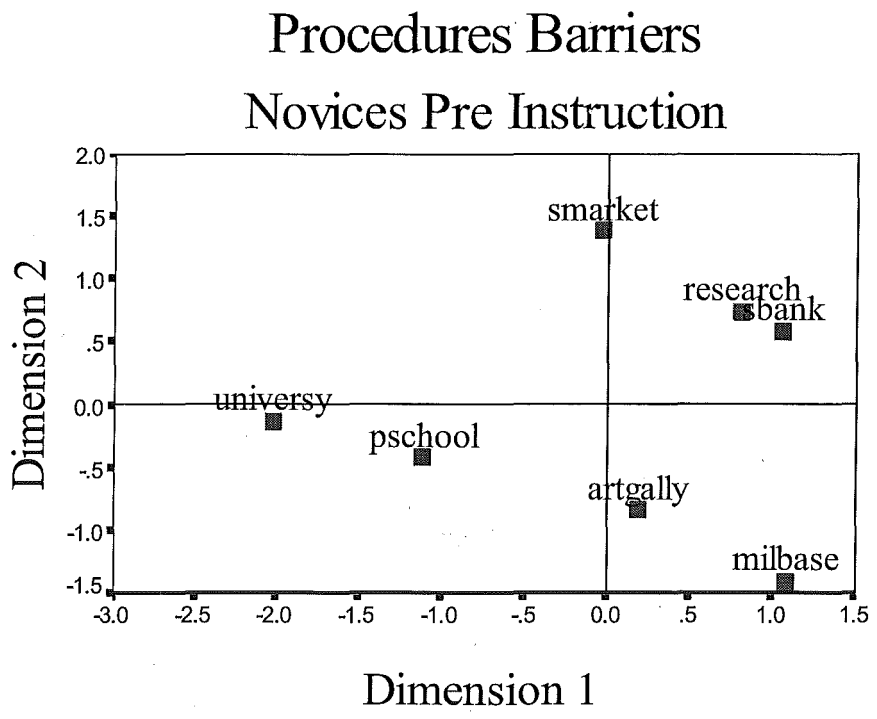


Figure 4: Novice Group knowledge structure for Procedural Barriers before Instruction.

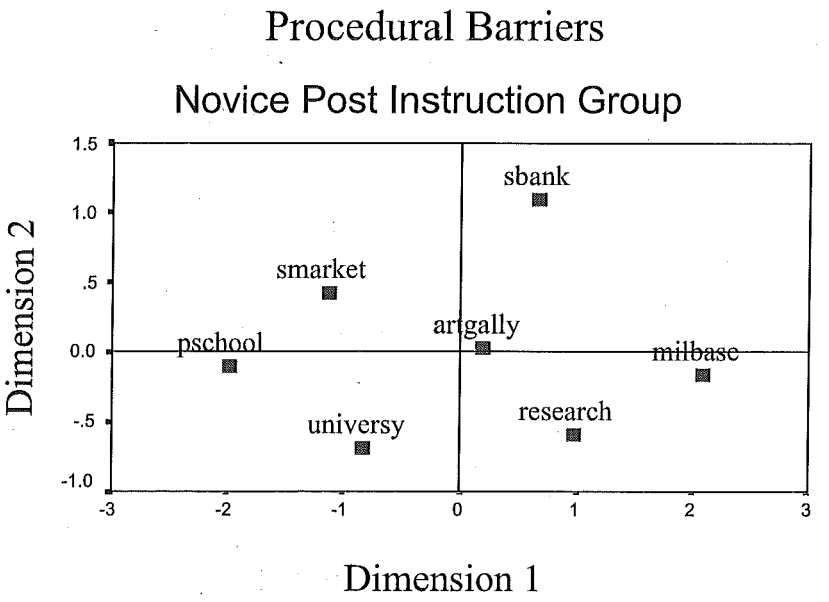


Figure 5: Novice Group knowledge structure for Procedural Barriers after instruction

The MDS statistical technique allows the comparison of knowledge structures through relative weightings on the defined dimensions of the knowledge structures. That is, the strengths of the individual knowledge structures can be displayed on the same dimensions (weighted averages). These weights of individual knowledge structures are represented as vectors in the 2-D conceptual space. The proximity of the knowledge structure vectors is a measure of the similarity of the knowledge structures. The Figure 6 shows the knowledge structure vectors for the Novice Group pre and post instruction for the Procedural Barriers, and the Expert group.

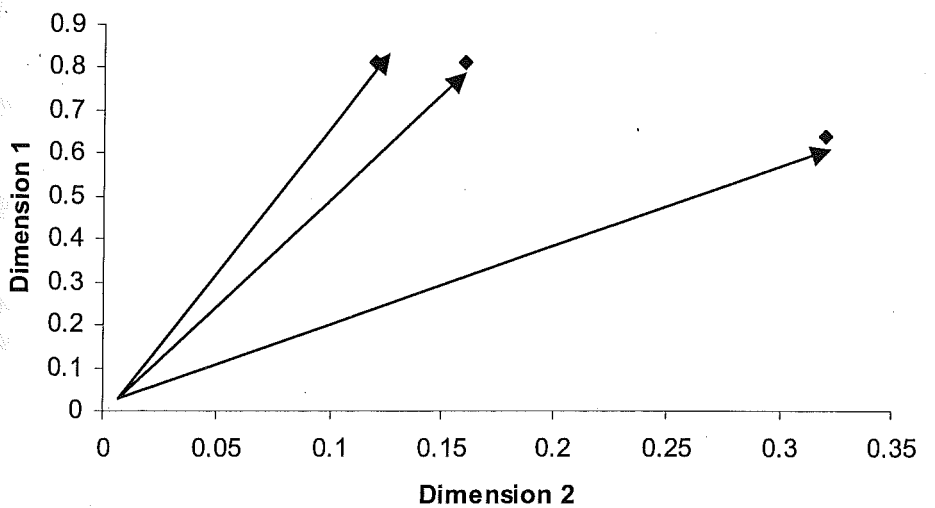


Figure 6: Comparison of the knowledge structures for the Novice Group (pre- and post-) and Expert Group for Procedural Barriers.

The learner-instructor model (Smith, 1986) proposes that the structure of concepts in the knowledge domain become more like the expert with instruction. In the Figure 6, the novice group pre instruction structure for procedural barriers is represented by the left vector according to Dimension 1 and

Dimension 2. Similarly, the central vector represents the novice group post instruction structure. The expert group's structure for procedural barriers is represented by the right vector. The vectors for the pre and post instruction for the novice groups show that the structure becomes more like that of the expert group with instruction.

CONCLUSION

The paper has described a novel approach to the understanding of novice learners in the knowledge domain of the *defence in depth* principle. These knowledge structures of students' perceptions of concepts in a *defence in depth* strategy have been compared to those of the experts in the knowledge domain and can be interpreted as the development of these structures with instruction. The novice learner group has displayed sufficient differences in 2-D knowledge structures to the experts' knowledge structure to claim incomplete understanding by the novices at this stage. However, with appropriate instruction and accumulated maturation of the learners with the *defence in depth* strategy, then it may be expected that these knowledge structures will tend towards the knowledge structure of an expert.

The proposed approach to a better understanding of the *defence in depth* principle by learners is speculative, and rigorous longitudinal studies of the change in knowledge structures for learners in this knowledge domain and other related domains is required to seek validation of the process. Provided that knowledge structures can be shown to be repeatable for individuals, then the reliability of the proposition that the knowledge structures of novice learners and experts will become more alike with suitable instruction is enhanced. The validity of the technique will be tested by further studies into the structure of knowledge of the *defence in depth* principle.

REFERENCES

Ausubel, D.P. (1966) Meaningful reception learning and the acquisition of concepts. In H.J. Klausmeier and C.W. Harris (Eds.) *Analysis of Concept Learning*. Academic press, New York.

Bruner, J.S., Goodnow, J., and Austin, G. (1956) *A Study of Thinking*. Wiley, New York.

Gonzalvo, P. et al., (1994) Structural representations in knowledge acquisition. *Journal of Educational Psychology*, 86(4), pp.601-616.

Koubek, R.J. and Mountjoy, D.N. (1991). Towards a model of knowledge structure and a comparative analysis of knowledge structure measurement techniques.

McGaghie, W.C., et al. (1998) Multidimensional scaling assessment of medical and veterinary student knowledge organisation of pulmonary physiology concepts. *Paper presented at the Annual Meeting of the American Educational Research Association*, 17p.

Michon, J.A. (1972) Multidimensional and hierarchical analysis of progress in learning. *In L.W. Gregg (Ed.), Cognition in Learning and Memory*. Wiley, New York.

Sireci, S.G. and Geisinger, K.F., (1995) Using subject-matter experts to assess content representation: An MDS Analysis. *Applied Psychological Measurement*, 19(3), pp.241-255.

Smith, C.L. (1986) An interpretation of the underlying dimensions of a knowledge structure of astronomy concepts for experts and novice learners. *In Nagy,P., The Representation of Cognitive Structures*, O.I.S.E., Toronto, pp.37-52.

Smith, C.L. and Robinson, M. (1999) The understanding of security technology and its applications. *Proceedings of IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology*, pp.26-37. Keynote Address. Madrid, Spain.

Steinberg, W.J. (1990) Differences between novice and expert knowledge structure, pre- and post-training, in a statistics and test theory domain. *Paper presented at the Annual Meeting of the Northeastern Educational Research Association*, 27p.

Streveler, R. (1994) Using multidimensional scaling to measure conceptual change. *Paper presented at the Annual Meeting of the American Educational Research Association*, 29p.

Takane, Y., Young, F.W. and de Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: an alternative least square method with optional scaling features. *Psychometrika*, 42, pp.7-67.

Public Street Surveillance: A psychometric study on the perceived social risk.

D. Brooks, C.L. Smith

Security and Applied Technology Research (SATR) Group
School of Engineering and Mathematics
Edith Cowan University, Western Australia
david.brooks1@defence.gov.au
clifton.smith@ecu.edu.au

Machine Vision Research Group
Department of Electrical and Electronic Engineering
Nottingham Trent University, England

ABSTRACT

This study quantitatively measured the social risk perception of public street surveillance, using spatial representation and multidimensional scaling (MDS). It utilized the psychometric paradigm, a method that presents risk perception in a two factor representation, being dread and familiarity to risk. The investigation showed the social risk perception of public street surveillance as low dread and familiar risk. MDS underlying dimensions presented public street surveillance as a low sense of risk perception and a low perceived community exposure to risk. This demonstrated a perceived social benefit, outweighing the perceived risk.

Key Words: *CCTV, Public Street Surveillance, Psychometric Paradigm, Risk Perception*

INTRODUCTION

Public street surveillance has grown enormously, is becoming common place and increasingly popular within society (Maley, 2000; Short & Ditton, 1998). Research shows that public street surveillance provides a decrease in levels of crime (Adam, 1998; Horne, 1998), but research has also shown that this may only be for a short period of time and only in certain crime categories (Ditton, 1999; Short & Ditton, 1998; Waters, 1996; Brown, 1995; Tilley, 1993).

Public street surveillance is often portrayed as the all purpose security tool that will greatly enhanced the level of protection of personnel and asset against risk. The security industry is quick to show the high performance of public street surveillance, with a typical example being "CCTV continues to be the buzz word around the country, most councils look to the Brisbane experience for arguments to convince ratepayers of the importance of the gadgetry." (Adam, 1998, p.30). It can be argued that the majority of public street surveillance media coverage is of a positive nature, with little or virtually no adverse media coverage, further promoting the introduction of public street surveillance into society.

Due to this increase, there was a requirement to measure the perceived social risk of public street surveillance. This paper will describe a recent research study that quantitatively measured this social risk perception. Closed Circuit Television (CCTV) has a wide and diverse application, with public street surveillance occupying a domain of CCTV. This research defined public street surveillance as a CCTV system that is located within, or is able to view, a public place.

BACKGROUND

There are numerous issues that could alter the social risk perception of public street surveillance. It can be argued that it will not be a single incident or that it will be immediate, but that changes will be slight and extend over a period of time. Issues that could change the risk perception of public street surveillance include the:

- Professionalism of the security industry, as a concern raised by Tate (1997) and how the industry manages, operates and promotes public street surveillance;
- Type and extent of media coverage;
- Perceived and applied effectiveness of legislation;
- Level of understanding individuals, groups and communities have of public street surveillance;
- Level of protection public street surveillance actually provides or is perceived to provide; and
- Community concern over civil and privacy issues.

Currently, the general perception of public street surveillance is that it improves safety and is not a social concern. In Dundee (UK), after public street surveillance was installed in the city center, a survey found that 96% of those surveyed through that public street surveillance would not infringe on civil liberties (Horne, 1998, p.321). More recent research by Ditton found that "67% ... did not mind being observed by street cameras."(1999). The Australian Privacy Commission stated "Queries about the legality of video surveillance were also common", being rated 6.9% of all enquires outside their jurisdiction (1998, p.43). But generally, it appears that the public views street surveillance as a benefit and therefore, an acceptable social risk.

It can be argued that with increasing exposure to public street surveillance and a growing public awareness, this view may alter. Slovic, Fischhoff and Lichtenstein reinforced this, when they stated that the "frequent discovery of new hazards and the widespread publicity they receive is causing more and more individuals to see themselves as the victims, rather than the beneficiaries, of technology." (1986, p.3). A number of authors have raised concern over the social affect of surveillance (Ditton, 1999; Davies, 1998; Tomkins, 1998; Waters, 1996). As Thompson discussed, "an individuals perception of ... risk can change. So can the level of risk that they are prepared to accept. These changes, which can be large, sudden and widely spread" (1982, p.62). The society we live in defines its own level of risk, not the expert or industry.

The purpose of this study was to assess the social risk perception of public street surveillance, with the research outcome being to demonstrate:

- A theoretical model to measure the social risk perception of public street surveillance;
- The social risk perception of public street surveillance;
- Whether public street surveillance is currently a sociably acceptable risk; and
- Whether public street surveillance significantly affects risk perception of certain demographic groups.

THEORETICAL FRAMEWORK

A suitable model to assess the level of risk perception was the psychometric paradigm. The psychometric paradigm is a method that attempts to assess and understand risk perception, and therefore risk acceptance to activities and technologies. As Slovic stated "psychometric paradigm, which uses psychophysical scaling and multivariate analysis techniques to produce quantitative representations or cognitive maps of risk attitudes and perceptions." (1987, p.281). This results in a two factor analytical representation (Figure 1), with the factor one axis being defined as *dread risk* and factor two axis being defined as *familiar risk.*

Dread risk is a dominating risk factor and can be made of various characteristics (Table 1), which were found to be highly correlated (Slovic, 1992, p.121). The other factor is *familiar risk*, again made

up of various characteristics (Table 1). This investigation (α=0.7240) tested the two risk factors of *dread risk* (r=0.2290) and *familiar risk* (r=0.2578).

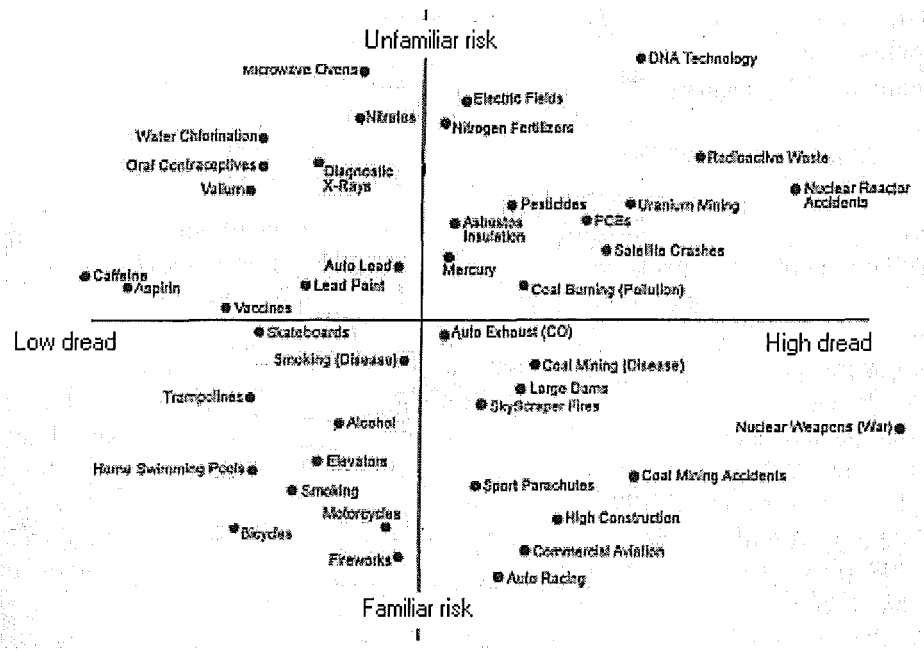


Figure 1. Psychometric Paradigm: Location of hazards.
(Adjusted from Slovic, 1997, p.236)

Table 1 - Characteristics of risk: Independent variables.

FACTOR 1 – *DREAD RISK*		FACTOR 2 – *FAMILIAR RISK*	
Low dread	High dread	Familiar	Unfamiliar
Controllable	Uncontrollable	Observable	Not observable
Low risk to future generations	High risk to future generations	Known to those exposed	Unknown to those exposed
Voluntary	Involuntary	Old risk	New risk
		Effect immediate	Effect delayed

The measure of each factor defined the perceived level of risk towards certain activities or technologies. A high sense of *unfamiliarity* and *dread* may lead to an increase in the perception of risk. Whereas *familiarity* and a low sense of *dread*, may lead to a reduced perception of risk. As the two factors alter within the community, so will the level of perceived risk for certain activities or technologies felt within that community. This paper used the term *risk perception*, cognizant that it takes the social psychology definition and includes attitudes, beliefs and judgements.

MULTIDIMENSIONAL SCALING

Multidimensional Scaling (MDS) is a statistical technique within the area of multivariate data analysis. MDS reduces complex dimensional data and presents these data in a spatial representation. "This ability to spatially represent a complex set of data is the major feature of MDS" (Smith, 1984, p.89). Cox and Cox define MDS as "the search for a low dimensional space, usually Euclidean, in which points in the space represents the objects, ... such that the distance between the points in space, ... match, as well as possible, the original dissimilarities." (2000, p.1).

The reduction in data complexity, through presentation in dimensional space, allows hidden structure to be shown in data. This demonstrates object proximity, with *proximity* being how similar or dissimilar objects are or are perceived to be (Kruskal & Wish. 1978). MDS was developed and theorized within behavioral science, but is now used by many disciplines (Cox & Cox, 2000), such as psychology to interpret perception, sociology to determine structure of groups, anthropology to compare different cultural groups, economist to determine consumer reaction to goods, and educators to research the structure of intelligence (Smith, 1984, p.88).

The process of MDS begins with a set of objects, being the activities and technologies, with measured dissimilarities between pairs of objects and placement within a half matrix format. A configuration of points is sought in dimensional space, with each point representing an object. The aim of MDS is to find a dimensional space configuration where the points distance *match* as close as possible, the paired dissimilarities. The different notions of *matching* defines the different techniques of MDS (Cox & Cox, 2000). MDS was used within the study as an additional statistical procedure to elicit further underlying risk perceptions from the research data. Primary models used were the Classical Euclidean Distance and Individual Differences (Weighted) Scaling (INDSCAL) models.

Classical scaling treats dissimilarities as Euclidean distances. Young and Householder, in the 1930s, demonstrated that a matrix of distances between points in Euclidean space can be preserved from point coordinates (cited in Cox & Cox, 2000). This was further developed and made popular by Torgerson (1952). The Euclidean distance algorithm is:

$$\delta_{re} = \left\{ \sum_{i} (xri - xsi)^2 \right\}^{1/2}$$

While the INDSCAL model, developed by Carrol and Chang (1970), coverts dissimilar data into distance estimates. Weightings are found by least squares and individual distances are doubled centered to produce matrices. Recurring least square is then applied, until convergence is achieved (Cox & Cox, 2000), with the INDSCAL algorithm being:

$$\delta_{rs} = \left\{ \sum_{i} wi\,(xri - xsi)^2 \right\}$$

METHODOLOGY

The investigation developed a seven point Likert risk perception survey questionnaire, containing 60 questions. Three additional questions included the participants gender, age group and distance to the center of the geographical research nucleus. The survey questionnaire contained the five activities and technologies of public street surveillance, radioactive waste, home swimming pools, drinking water chlorination and coal mining disease.

The additional four activities and technologies not only provide spatial relationship comparisons of where public street surveillance was located, but also allowed a comparison of previous psychometric studies (Bouyer, Bagdassarian, Chaabanne and Mullet, 2001; Slovic, 1997). They were chosen, as according to Slovic (1997), they represent one object from each quadrant of the spatial factor representation model.

STUDY TARGET POPULATION

The target population (N=2106) were community members who lived and/or worked within 0.5km distance of Rockingham beachfront, within the City of Rockingham, Western Australia, and were ≥16 years old at the time of survey. The sample participants (N=169) were random volunteers selected from the target population. At the geographical nucleus of the target population, the shire proposed

the installation of a public street surveillance system. The study was completed before this system was installed, with the intent to complete a post study six-months after the system had been installed.

The research demographics comprised of females (49.1%, N=83) and males (50.3%, N=85), with the majority group being those aged 56 years old or greater (31.34%, N=53). This was followed by those aged between 36 to 45 years old (24.3%, N=41) and 46 to 55 years old (22.5%, N=38). The majority of the participants (86.4%, N=146) worked and/or lived ≤4km from Rockingham beachfront.

RESULTS AND INTERPRETATION

The survey data were analyzed and interpretations resulted in some significant findings. This included the measured risk perception of public street surveillance, lower female risk perception, dominant risk perception characteristics and the *underlying* MDS dimensions of public street surveillance.

RISK PERCEPTION OF PUBLIC STREET SURVEILLANCE

Public street surveillance (mean total of all participants) occupied the same spatial quadrant as home swimming pools, with low dread and familiar risk. The spatial factor representation is shown in figure 2. The population perceived public street surveillance as a low risk to future generations, that exposure was voluntary, the risk was observable, that they understood the risk, that the risks are known and would be immediate. This results in public street surveillance having a perceived low social risk perception, demonstrating a social benefit that outweighed the perceived risk to the community.

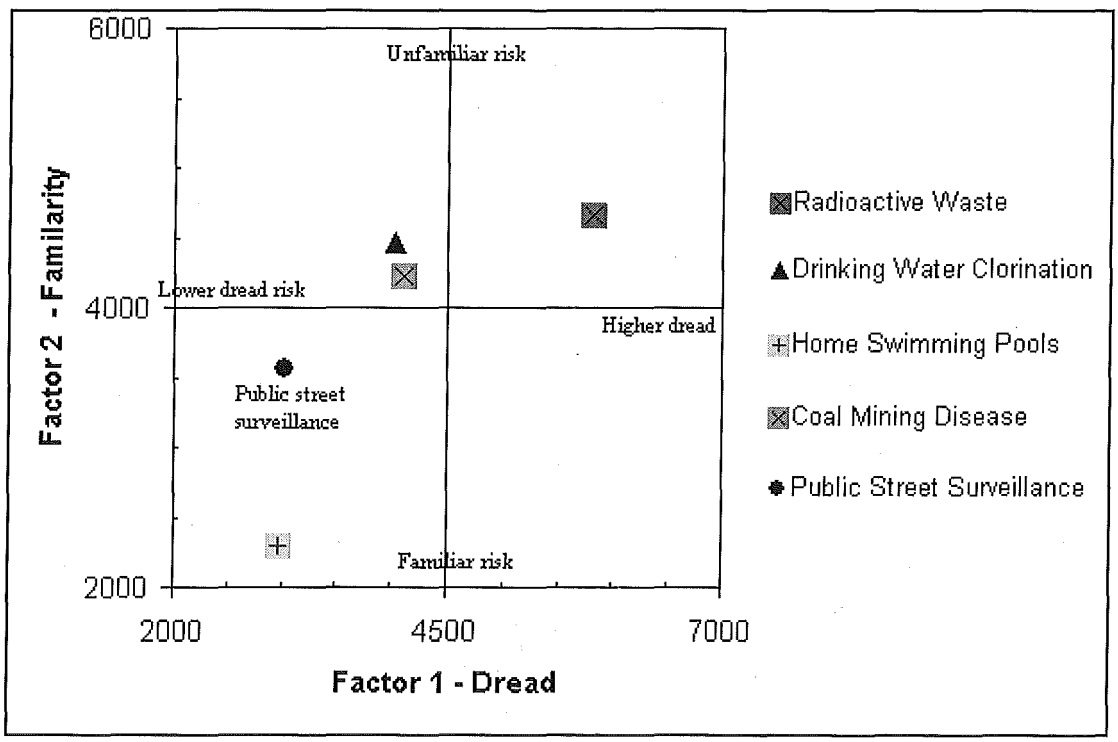


Figure 2. Spatial factor representation of all activities and technologies.

GENDER RISK PERCEPTION

Females exhibited greater dread and/or unfamiliarity levels across all technologies and activities, apart from public street surveillance. This had been demonstrated by a number of previous studies (Bouyer,

Bagdassarian, Chaabanne and Mullet, 2001; Slovic, 1992). As Slovic stated "sex differences were quite interesting. Close to two dozen studies have found that women have a higher perceived risk ... than men." (1992, p.129).

But females demonstrated a lower level of dread and felt more familiar with public street surveillance. This appeared to indicate that females not only feel safer when public street surveillance is present, but that they feel that public street surveillance is a social benefit which outweigh perceived risk. Therefore, females would support the introduction and maintenance of these systems.

AGE & DISTANCE RISK PERCEPTION

For public street surveillance the sense of risk perception did not appear to be related to age, which produced no significant difference and the closest spatial cluster of all activities and technologies. Also, the distance that the participants lived and/or worked from the nucleus of the public street surveillance system did not cause any significant different in the sense of risk perception. Again, this factor presented the closest dimensional cluster of all tested activities and technologies.

PUBLIC STREET SURVEILLANCE CONTROLLABILITY

The characteristics of each activity or technology were analyzed, showing how each was related. As Slovic stated, "every hazard had a unique pattern of qualities that appeared to be related to its perceived risk." (1992, p.121). Figure 3 presents the mean profiles for each item characteristic.

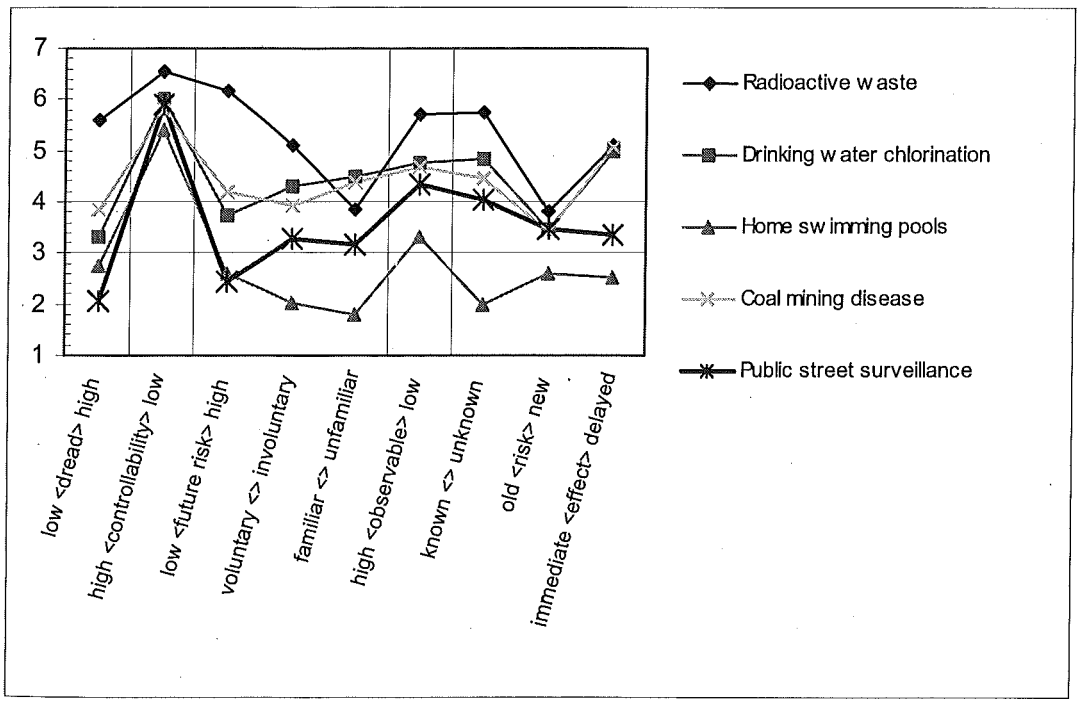


Figure 3 Factor characteristic profiles of each activity and technology.

Public street surveillance produced the lowest level of *dread* and *risk to future generations*. But the *familiarity* characteristic, public street surveillance was located close to the center. This may indicate neutral risk perception or that little social thought had been given to public street surveillance.
Public street surveillance controllability produced a significantly lower result then radioactive waste [$t(163)=4.3969, p=0.000$], but significantly greater result then home swimming pools [$t(163)=-2.6651, p=0.0085$]. Extracting the characteristic of control, located public street surveillance towards a higher dread, but still within the familiar quadrant. The difference between total dread and control

was significant [t(165)=21.7,p=0.0000]. It appeared that the community had a social concern over the ability to ensure appropriate public street surveillance control. But these representations did not demonstrate *underlying* individual differences with the participants, or represent concept space. Therefore, the data were further analyzed using MDS.

MDS UNDERLYING RISK PERCEPTION OF PUBLIC STREET SURVEILLANCE

Using the Euclidean distance model, the total means of all activities and technologies were tested. This presented the spatial representation distribution [interval data, stress 0.0039, RSQ 0.9999]. This demonstrated the group space of the activities and technologies, and formed the *underlying* configuration of points (Cox & Cox, 2000). Placed onto the group space, was the psychometric spatial factor representation of factor *dread* and *familiarity*.

The two MDS dimensions were reduced to one-dimensional representations. This produced a dimension 1 spectrum that matched in distribution, both *dread* and *familiarity*. Drinking water chlorination and coal mining disease *dread* were in reverse order, but this reversal was not considered significant due to the spatial similarity of these objects. Therefore, this dimension appeared to measure the "community risk perception" of the activities and technologies.

Dimension 2 produced a spectrum that placed radioactive waste and home swimming pools together at one end of the distribution. This showed that the spatial representation (Figure 2) did not fully demonstrate the *underlying* risk perceptions. Dimensions 2 was defined as the "perceived community exposure to the risk".

This placed public street surveillance within its "own" quadrant and indicated that it had unique characteristics, when compared to the other activities and technologies. Applying the MDS dimensional characteristics (Figure 4), public street surveillance appeared to have a low level of risk perception and a low perceived community exposure to risk.

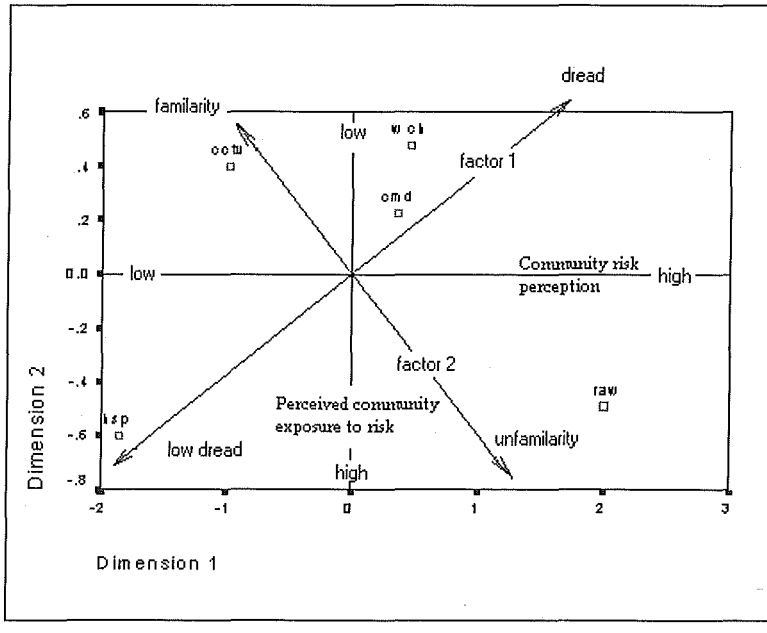


Figure 4. MDS spatial representation of concept space for all activities and technologies.
(raw= radioactive waste, wch= drinking water chlorination, hsp= home swimming pools, cmd= coal mining disease, cctv= public street surveillance)

MDS INDSCAL

INDSCAL was applied to the sub group of gender for all activities and technologies. A spatial representation distribution [interval data, stress 0.0789, RSQ 0.9713, matrices mean] was presented, which appeared similar to the spatial factor representation (Figure 2).
Unlike the MDS spatial representation (Figure 4), the gender spatial distribution dimensions matched the factor distribution. Dimension 1 was similar to *dread* and dimension 2 was similar to *familiarity*. But this distribution placed drinking water chlorination and coal mining disease in the quadrant, higher *dread* and *unfamiliar risk*. This plot opposed the spatial factor representation distribution (Figure 2), by shifting the axis location of factor 2 *familiarity* further to the left.

INDSCAL has demonstrated that public street surveillance had a similar level of *dread* as home swimming pools. This also demonstrated that public street surveillance was located within the spatial factor representation (Figure 2), quadrant of low dread and familiar risk. Public street surveillance was located close to the center of dimension 2 *familiarity*, showing a neutral response and that there may have been little thought given to the social issues of public street surveillance. This supports the factor characteristic profiles (Figure 3).

STUDY LIMITATIONS

This study tested a relatively low number of activities and technologies, when compared to other psychometric studies. Using the MDS technique, this low number of activities and technologies reduced the ability for the data to produce and provide defined clusters (Kruskal & Wish, 78. pp. 43-46) and increased the subjective interpretation of the dimensions. This requires further validation to quantify and confirm the MDS dimensions of "risk perception" and "perceived community exposure to risk".

The paper presents the findings from the pre study phase. The post study will attempt to measure change in the location of the activities or technologies. This will allow further study correlation and validation, and whether the community risk perception changed with increased exposure to public street surveillance.

OUTCOMES

This study has applied a theoretical model to measure the social risk perception of public street surveillance, utilising the psychometric paradigm. Public street surveillance was presented as occupying the social risk perception of low dread and familiar risk, being a relatively safe social position. This resulted in a low social risk perception and demonstrated a perceived social benefit, which outweigh the perceived risk to the community.

This outcome was supported by MDS analysis, which provided the additional *underlying* dimensions and presented public street surveillance within its own quadrant, indicating unique characteristics. Applying MDS, public street surveillance appeared to have a low sense of risk perception and a low perceived community exposure to risk. Therefore, the community will generally support the introduction and maintenance of public street surveillance.

For public street surveillance, females presented a lower social risk perception than males. This appeared to indicate that females feel safer when public street surveillance is present. But that there has been limited thought given to the social issues of public street surveillance. It appeared that the community had a social concern over the ability to ensure appropriate public street surveillance control. This may demonstrate the underlying social concern of public street surveillance and why it occupies its own unique area within the MDS spatial representation (Figure 4), as once isolated, this risk characteristic was significantly higher than *dread risk*.

CONCLUSION

Public street surveillance appeared to occupy a safe and relatively non-adversary position within society, but risk perception is capable of change. With little thought given to CCTV and minimal adverse media coverage, this further promotes and maintains public street surveillance. Although this study isolated public street surveillance from CCTV, lay people will not and any social risk relocation will affect all domains of CCTV.

This appeared to be particularly valid over the control of CCTV and how that control is applied, managed and maintained. These issues are likely to be driven via media coverage over high profile incidents. Whether control is applied through legislation or self-regulation, it will have to be applied to alleviate both real and perceived social risks. But while the perceived social benefit outweigh the social risk, CCTV risk perception is unlikely to change.

REFERENCES

Adam. (1998, August). Security Australia Magazine. Volume 18, Issue 7. Reed Business Information.

Brown, B. (1995). CCTV in town centers: Three case studies. Crime Detection & Prevention Series Paper 68. Home Office, UK.

Bouyer, M., Bagdassarian, S., Chaabanne, S., & Mullet, E. (2001). Personality correlates of risk perception. Risk Analysis, Vol. 21, No. 3.

Carrol, J.D. & Chang, J.J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalisation of "Eckart-Young" decomposition. Psychometrika (35), 283-319.

Cox, T.F. & Cox, M.A.A. (2000). Multidimensional scaling. Monographs on statistics and applied probability: 88. Chapman & Hall/CRC.

Davies, S.G. (1998). The effectiveness of CCTV. The case against: CCTV should not be introduced. International Journal of Risk, Security and Crime Prevention, 377-331.

Ditton, J. (1999). The effect of closed circuit television cameras on record crime rates and public concern about crime in Glasgow. [on-line]. The Scottish Office Central Research Unit Main Findings, No. 30. Available: http://www.scotcrim.u-net.com/reserachc2.htm, [29, April 2002].

Federal Government. (1998). Tenth Annual Report on the Operation of the Privacy Act. [on-line]. Australian Privacy Commission. Available: http://www.austlii.edu.au, [27, April 2002].

Horne, CJ. (1998). The Effectiveness of CCTV has been proven: Should CCTV be introduced into all town and city centres. International Journal of Risk, Security and Crime Prevention, 317-326.

Kruskal, J.B. and Wish, M. (1978). Multidimensional scaling. Sage University series no. 07-011. Sage Publications, Beverly Hills, London.

Maley, P. (2000, August). CCTV: The unseen eye. GEN Magazine, 18-19.

Short, E. and Ditton, J. (1998). Seen and now heard: Talking to the targets of open street CCTV. British Journal of Criminology, Vol. 38, No. 3, Summer, 404-428.

Slovic, P. (1987). Perception of risk. Science, Vol. 236, 280-285.

Slovic, P. (1992). Perception of risk: Reflections on the psychometric paradigm. In S. Krimsky & D. Golding (Eds.), Social theories of risk. (pp.117-152). Praeger Publishers.

Slovic, P. (1997). Risk perception and trust. In V.Molak (Ed.), Fundamentals of risk analysis and risk management. (pp.233-245). CRC Press.

Slovic, P. Fischhoff, B. Lichtenstein, S. (1986). The psychometric study of risk perception. Risk Evaluation and Management. Plenum Press. New York.

Smith, C.L. (1984). Learning astronomy and the organization of astronomy concepts in semantic memory. Doctoral dissertation, Murdoch University, Perth, Western Australia.

Tate, P. (1997). Report on the security industry training. WA Department of Training. WA Government Publishing.

Thompson, M. (1982). A Three Dimensional Model. Essays in the Sociology of Perception. Routledge and Kegan Paul, London.

Tilley, N. (1993). Understanding public car parks, crime and CCTV: Evaluation lessons from safer cities. Police research group. Crime prevention unit series paper no. 42. UK Home Office.

Tomkins, P.L. (1998, April). Panopticon: Technology, the individual and social control on the early 21st century. The Journal of the Royal United Institute for Defence Studies.

Torgerson, W.S. (1952). Multidimensional scaling: 1. Theory and method. Psychometrika (17), 401-419.

Waters, N. (1996). Street surveillance and privacy. [on-line]. Australasian Legal Information Institute. Available: www.austlii/ii.edu.au, [27, April 2002].

Peer-to-Peers a Year After the Decline of Napster

Bengt Carlsson

Dept. of Software Engineering and Computer Science
Blekinge Institute of Technology, Sweden
Email: bengt.carlsson@bth.se

ABSTRACT

When Napster closed down a year ago an alteration occurred from the original equal peers to a situation where the user opens up the entire local disk to anybody without any control. Today different P2P tools are infected by adwares, spywares and/or trojans. This infection causes both a risk of vulnerability against security and privacy attacks and a poorer performance of the local computer system. Personal firewalls and anti spywares may be used to prevent these attacks. Computer systems with and without ad/spywares and/or personal firewalls were investigated. A positive correlation between ad/spywares installed and number of pop-ups were found. A firewall reduced but did not exclude the number of pop-ups. When both ad/spywares and P2P programs were activated the computer system became unstable.

Keywords: Peer-to-peer, spyware, adware, arms race.

INTRODUCTION

The two major concerns about peer-to-peer (P2P) systems are anonymity and non-censorship of documents. Basically a P2P system consists of a society of equal peers acting both as servers and clients. With increased bandwidth access and efficient file compression facilities it is now possible to share large files like MP3 and video files, violating the copyrights of record and movie companies.

Last spring (2001) Napster was the predominant MP3 file-sharing program with millions of users. The Recording Industry Association of America (RIAA) has since then highlighted copyright protection by forcing Napster to close down. There is an increased lack of equality within the P2P society with one group of peers supplying the tools and one group using the tools. This can be highlighted by the introduction of spywares and adwares.

Napster did not contain any ad/spywares at all. Instead of letting Napster become a standard within the P2P file sharing community, we are witnessing a transition towards P2P-tools where the users have to accept different ad/spyware within the program. The main focus presented here is the transition from the original equal peers to a situation where the user opens up parts of or the entire local disk to anybody without any control, i.e. the opposite to the original idea about file-sharing.

In section 2 a background is introduced to the concept of an antagonistic P2P society. In section 3 the current situation within P2P is presented followed by an investigation of the behavior of the current spywares in section 4. A discussion follows in section 5 and finally some conclusions are drawn.

BACKGROUND

The ongoing battle within P2P systems is neither a new phenomenon or something unexpected. The issue of security and privacy from a social and business point of view has been addressed by both computer scientists (Anderson 2001, Schneider 2000) and economists (Shapiro et al. 1999).
When analyzing very large groups like the actors on the Internet, it may be possible to find patterns for the behaviors of peers. The success of a distributed P2P system is dependent on both cooperating coalitions and antagonistic behaviors if there is a conflict of interest involved. Within and between the groups conflicting interests should be expected. The goals of a peer are usually provided by a human, often its owner.

Achieving these goals may involve humans acting in a competitive surrounding. We will use the following terms/concepts to describe these competitive activities:

Humans with *Machiavellian intelligence* (Dunbar 1997), i.e. bringing out self-interest at the expense of others. This is a manipulative activity directed against other individuals.

An *arms race* (Dawkins 1982) between individuals or between groups of individuals, i.e. the (antagonistic) activities made by one group are retorted by countermeasures by the other group, which in turn makes the first group react, and so on.

The *tragedy of the commons* (Hardin 1968), describes a situation where the costs caused by the action of a selfish individual are shared by all participants, while the selfish individual gets all benefits of this action. There is a risk that everyone will get worse off in a competitive surrounding.

From a general assumption of humans being selfish and acting as Machiavellian beings, an arms race is supposed to evolve. As a negative consequence all the peers may suffer from the activity of a single peer (the tragedy of the commons). A possible positive consequence is an evolving robustness for the P2P society against unforeseen malicious activities.

Adar et al. (2000) have shown the unwillingness for a majority of users to share files within a P2P system. More in general; an individual within the user group will benefit from cooperation if it is possible to identify other users and the cost for doing services towards other users are negligible. In practice it is sufficient to have a fraction of the users cooperating to maintain a well performing system. The ongoing battle within file-sharing tools and the conflict of interest among different actors could be seen as an *arms race* within peer-to-peer tools. Within the P2P field the *tragedy of the commons* metaphor is used by several researchers (e.g. (Adar et al. 2000; Oram 2001) for explaining overexploitation of the resources.

In figure 1 a sketch of the different actors involved within a P2P society is shown. A peer, or end user, downloads a P2P program and gets additional programs as a disadvantage. Most P2P tools today are given away for free so additional programs provide the main earnings for the P2P program developer. Additional programs make their earnings mainly from selling commercials. Unlike radio or TV commercials a computer-based advertisement may be personally addressed. A personal firewall keeps track on ingoing and outgoing network traffic. Together with so-called anti-spywares it is possibly to find and remove undesired components. In section 3 these different components are described more in detail.

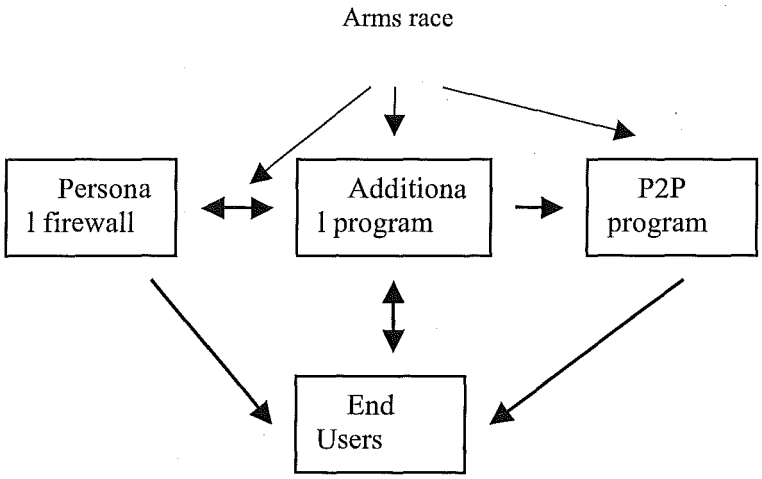


Figure 1 The information flow involved using P2P tools

An arms race occurs in at least three different occasions:

Between different P2P tool and between a P2P tool and an outside actor like RIAA. A successful P2P tool avoids being sued by RIAA and attracts a lot of end users. This subject is further analysed in Carlsson (2001) and will not be a main scope of this paper.

Between additional programs and between different advertisers using these programs. A successful advertiser both reaches its group of consumers and prevents others from doing so. Examples of such behaviours are shown in section 3.

Between additional programs and personal firewalls together with anti spyware programs. This arms race is similar to the race between virus and anti-virus programs with no obvious finite winner. We will come back to this issue in the discussion part.

From an end users point of view s/he wants to download files from other peers. S/he probably accept others to check and download files locally installed, and may even accept advertising pop-up windows now and then. What is more important, the end user is probably not aware of all watching spywares and foistwares (see section 3) bundled together with the installed P2P program. In section 4 we investigate the performance of the P2P system from the end users point of view and in section 5 we discuss a more general privacy threat.

PEER-TO-PEER TOOLS

In the spring of 2001 the fight between the Recording Industry Association of America (RIAA) and Napster had just begun. At that time three different P2P tools were investigated; Napster, BearShare and MusicCity (Carlsson 2001). Napster used a central file register while BearShare, a Gnutella P2P-tool, was equally distributed among the different peers. MusicCity was in between, it did not maintain a central content index but required centralized user registration and logon.

Compared to Napster, both MusicCity and BearShare contained only a small fraction of users. These users could download more MP3 files, because Napster tried to filter out copyright protected files. The Napster alternatives MusicCity and especially BearShare consumed more resources both by default and by having a higher rate of aborted file transfers.

When former Napster users are leaving for other P2P tools, this causes higher bandwidth usage for these users and thus increases the communication needs over the Internet (unlike the centralized Napster distributed P2P, e.g. BearShare, means heavily increased network loads). In Napster there is still some incentive for being cooperative because such a user is not anonymous. In BearShare and MusicCity other users are more anonymous because one MP3 file may be downloaded from multiple sources.

The new P2P tools increasingly include features not desired by the users. These additional programs may be regarded as trojan horses because the main business purpose is to offer a P2P program for free and earn money from the additional programs. The additional programs are roughly divided into:

adware - applications generating commercial advertising, often as pop-up windows when surfing the net.
spyware - a program that secretly collects information about the user and sends this information back to the spyware-owner.

foistware - applications that secretly add and hide components within a local system. This is spywares behaving like viruses because new files may be generated and old files may be changed without the knowledge of the user. In the remaining of the paper we will not separate spywares and foistwares.

In the next section the increasing usage of advertising banners and spywares is investigated. The following programs, found in the investigation, may clarify how these different additional programs are working:

WhenUcom (SaveNow.exe) - distributes, as popup-windows, commercials based on visited web-places. Information about sex, age, address and e-mail is collected and distributed to other interested parties.

web3000 (msbb.exe) - collects information about the internet habits of the users. This program is hard to delete because it replaces a Windows system file with its own file.

GAIN (CMESys.exe and GMT.exe) - collects information about visited web pages, system settings, local software and personal information.

Many more examples could be found at http://www.cexx.org. Examples of the ongoing arms race are home page hijackers and dialler programs that have started to infect software. The function of the hijacker is to change the browser's default start page to point to their site. This site is almost always loaded with adwares. In a dialler program a dialler changes the ISP dial-out number to a high-charge per-minute phone number. Diallers may also silently disconnect and reconnect modems even when surfing the Web, as soon as the program is executed.

Today Napster has closed down, replaced by a group of other programs. MusicCity has changed location and name (Morpheus and Kazaa) and now belongs to the Gnutella clones, just like, as before, BearShare. Other upcoming programs like AudioGalaxy attract a lot of users (but is now in September 2002 closed down). This process is not just an exchange but an altering of the previous conditions, and that is why the concept of arms race may be a proper description.

INVESTIGATING CURRENT P2P TOOLS

For the investigation three different P2P tools with documented adware/spywares were chosen: Kazaa, Bearshare and AudioGalaxy. In all there is about 15 different additional programs reported for these programs and we were able to find the four processes mentioned above in the list of active processes. We did not try to find inactive processes so it is not possible to tell if all reported additional programs

are present. Kazaa for instance may upgrade automatically adding/deleting the current set of additional programs.

For each computer two out of three P2P programs were selected. The computers were used as close to normal as possible, reflecting the habits of a typical student with access to a fast Internet connection. The computers were shut down and restarted approximately every second hour.

Three different computers (Pentium 133MHz and Pentium 4 1,4 GHz) with standard programs installed (MS Windows XP, MS Office XP, ICQ, Eudora, Winamp and Norton Antivirus) were used. During approximately ten hours for each computer each of the following steps were performed:

NoPrograms - no P2P or additional programs were installed on the computers.
AddsOn/ZoneAlarm - P2P programs were installed but not run. Any additional program installed during P2P installation may run independently. ZoneAlarm, a personal firewall, is alert.
AddsOn - same as 2 but without ZoneAlarm alert.
AddsOn/P2P/ZoneAlarm - P2P and AddsOn are running and ZoneAlarm is alert.
AddsOn/P2P - P2P and AddsOn are running but ZoneAlarm is not alert.

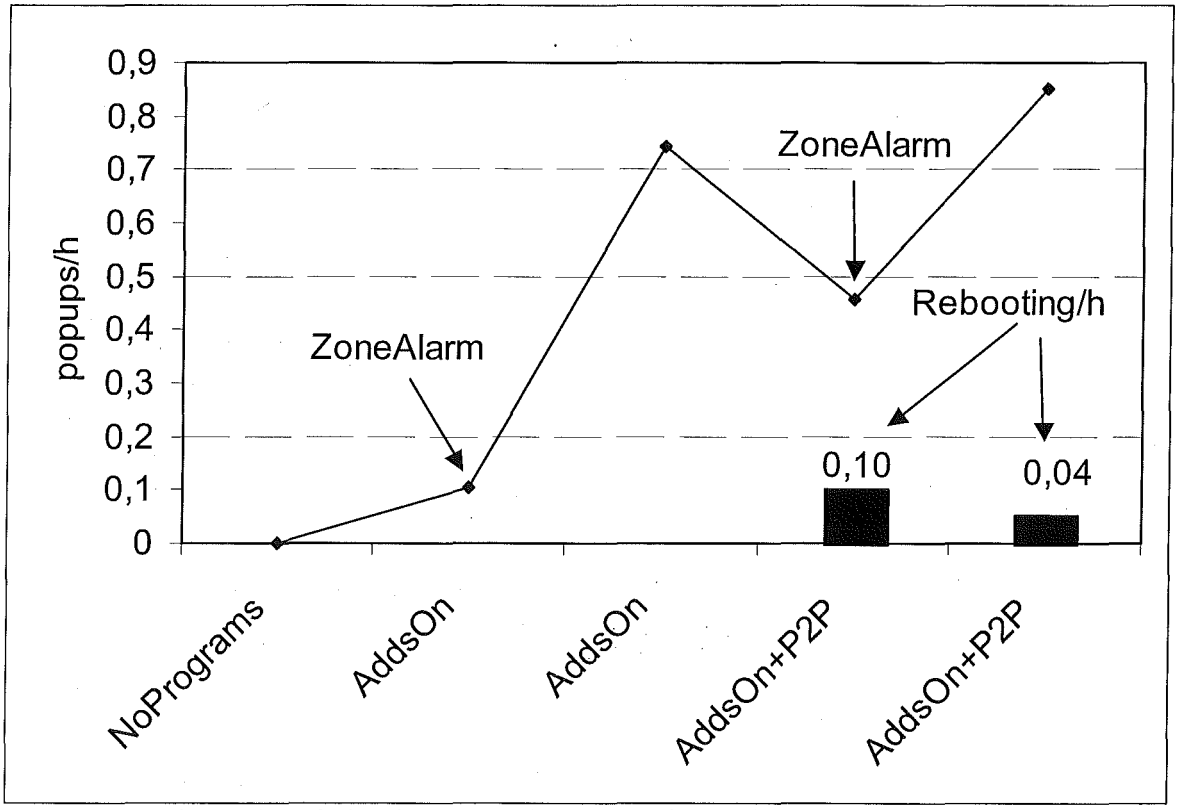


Figure 2. The number of advertisements and system crashes per hour when different programs are alert.

SYSTEM USAGE

Additional programs installed themselves more or less secretly as processes. When measuring with and without ZoneAlarm and with and without P2P tools active we found only small variations in activity. Most of the time between 30 and 40 MB of the primary memory was used for the additional programs. The CPU and network usage varied a lot depending on if the additional programs were activated or not. At most SaveNow used 38% of the processor capacity (P4 computer) and together with a P2P program the additional programs used all available network capacity at several occasions. Most of the time the additional programs were resting.

What is more important, the computer system seems to be more unstable with additional and P2P programs installed. The test period is too small to draw any general conclusions but even an indication of system crashes on average every 10th hour with ZoneAlarm installed is conspicuous (see figure 2). One explanation for the increased instability is the repeated attempts from additional program to reach the net after being stopped by ZoneAlarm.

POP-UPS

Pop-ups appear when P2P and additional programs are installed. Figure 2 shows that some pop-ups occur without any P2P programs being started and even with a personal firewall active. Most pop-ups occur with active P2P and passive firewall.

Pop-ups generated by visited web pages instead of the installed programs may explain part of the pop-up frequency above. About 2/3 of all pop-ups were generated by the installed programs. The rest of them may also be generated because original pop-ups are exchanged or a new one is generated visiting a certain web place.

DISCUSSION

Most programs used together with P2P tools contain adware, spywares or trojans. The differences between these programs and traditional virus and trojans are mainly how they are treated by different anti-virus programs. If an additional program is classified and defeated as a virus it will probably disappear because most users have active anti-virus programs installed. If instead a program is classified as a spyware, most users do not have anti-spywares or local firewalls installed. The legitimacy is confirmed by including the use of a spyware program in the licence agreement and/or having a yes/no option for installing it. In practice additional programs are often integrated with the P2P part so most users are forced to install spywares in their computers if they want to use P2P tools.

Companies saying they do not misuse information is the only assurance against collected information. We do not for sure know what is collected, because information is not sent back in plain text. We neither know who else will get the information and for what purpose, companies regularly claim permission to forward information. In hardly any other area would we accept such an undefined guarantee, so why? The most obvious answer is that everyone else is doing the same. Even Microsoft collects information about (competing) programs during Windows installation. This makes it hard to construct a general anti spyware tool that manages every threat against the users privacy.

Another disadvantage is a poorer performance for the system. Just as junk or spam mails may be regarded as an additional cost because of a waste of time, so may additional pop-up advertisements be treated. The trend today is bigger pop-ups that are hard to close down and in many cases pop up again. Our investigation shows an increased number of pop-ups when using P2P tools, which even show up without using the programs. This is not like listening to commercial radio and watching TV stations where there is a choice if you want the commercials or not (you can always turn it off).

Another uncontrollable consequence of additional programs is an increased risk of an unstable system. This investigation does not prove but indicates such a risk. With people logged on for longer and longer periods of time this is a big problem. It is not you making a mistake but invisible programs causing a crash that causes the problem.

We are witnessing the tragedy of the commons within P2P systems. Very few participants gain from the current situation. In the short run maybe RIAA and other copyright protecting organizations welcome poorly working file sharing tools. In the long run almost everyone will lose if the reliance upon Internet based computer system fails. Personal firewalls may have the same role as anti-virus tools but with two weakened exceptions: they are hard to configure and it is hard to define if we are

dealing with a smart business program or an insidious intruder. The end user reaction will probably tip the scale. With increasing problems more adwares and spywares will be classified as viruses. If companies only receive bad reactions they will probably not pay for generating pop-ups. Many of the current problems may be classified as teething problems that hopefully disappear with a more mature P2P society.

CONCLUSIONS

Since Napster closed down P2P users have moved to programs infected with spywares and adwares. This transformation causes both a risk of vulnerability against security and privacy attacks and a poorer performance of the local computer system. Personal firewalls and anti spywares reduce but do not prevent the attacks. We are witnessing the tragedy of the commons within P2P systems. In the long run almost everyone will lose if the reliance upon Internet based computer system fails. Hopefully more efforts from both security/privacy protecting companies and end users will block the way for these attacks in the future.

Acknowledges: Thanks to Björn Folbert, Magnus Persson and Henrik Svensson for doing the experimental studies about P2P-tools. Besides the anonymous reviewers I also want to thank Per Mellstrand for comments and Martin Hylerstedt for proofreading.

REFERENCES

Adar, A., and Huberman, B.A., Free riding on Gnutella, FirstMonday peer-reviewed journal on the Internet *http://firstmonday.org/issues/issue5_10/adar/index.html,* 2000

Anderson, R., *Security Engineering - A guide to Building Dependable Distributed Systems*, John Wiley &Sons, Inc, New York, 2001

Carlsson, B., The Tragedy of the Commons - Arms Race within Peer-to-Peer Tools, in eds. Omicini, A., Petta, P., and Tolksdorf, R., *proceedings of the 2nd International Workshop Engineering Societies in the Agents' World, Lecture Notes in Artificial Intelligence 2203, Springer-Verlag*, 2001

Dawkins, R., *The Extended Phenotype*, W. H. Freeman and Company, Oxford, 1982

Dunbar, R., *Grooming, Gossip and the Evolution of Language*, Faber and Faber, 1997

Hardin, G., The tragedy of the commons, *Science vol. 162 pp. 1243-1248,* 1968

Oram, A., ed., *Peer-to-peer Harnessing the Power of Disruptive Technologies,* O'Reilly, Sepastopol, CA, 2001

Schneider, B., *Secrets and Lies: Digital Security in a Networked World*, John Wiley &Sons, Inc, New York, 2000

Shapiro, C., and Varian H.R., *Information Rules - A strategic Guide to the Network Economy*, Harward Business School Press, Boston, MA, 1999

Information Systems Competence: An Information Warfare Disaster Waiting to Happen

Edmond La Vertu

School of Management Information Systems
Edith Cowan University, Australia
E-mail: lavertu@vianet.net.au

ABSTRACT

Philosophically speaking, the IS industry is limited to the competency skill level when greater skill levels may be needed to effectively counter an information attack. Here is discussed the Dreyfus Skill Acquisition model and here it is applied to Information Systems (IS), and it is concluded that IS is in need of a rethink about what level of trained personnel it needs to counter a Systems Attack. It is posited that the IS industry has limited the training to 'rule following' or calculative rationality when it truly needs training in deliberative rationality, that is, intuitive problem solving as well.

Keywords: Information Systems, Information Warfare, Phenomenology of Skill Acquisition

INTRODUCTION

Here we consider a single Information Systems (IS) problem. This problem is that while technology changes at an almost dazzling pace, the ability of the commercial world to adjust its skills development above and beyond competent skill levels, is in my opinion, non-existent. IS technologists are, for the most part, not seen as an integral and multi-skilled part of the organisation, rather they are seen and treated as a separate, overhead creating, and skill limited bureaucratic function necessary for the operation of their business.

Years ago when I first entered Data Processing (DP) as it was then known, problems arose concerning costs, time to first billing, skills acquisition and utilisation of those skills, client and user security, and lifetime of the product before modification, and all of these and more, are today's problems as well. Also, all of these problems are people sourced yet it seems that now, as then, a rule based technological solution is demanded without consideration of the people concerned, their varied skills and skill levels, their culture and working environment, and their technological and other education.

Also, every piece of equipment used in any network is 'open'. That is, the knowledge about any Information Systems' most intimate workings is freely available from any reputable IT training school, public and private, or from the manuals supplied with the equipment or both. There are no secrets in commercially available IS hardware or software systems. A technical problem in one is a problem for all that use the same IS tools.

These are our real givens in an industry which, it seems to me, has yet to reach a maturity level where skills beyond competence exist and are readily accessible for problem resolution, training, planning and policy direction.

The conclusion reached in this paper is for me, surprising. Unless we begin a strong evolutionary movement change in the IS industry, these problems will remain, irrespective of the desires of the industry, the user or the competent IS industry technologist.

LITERATURE REVIEW

In seeking explanations to this problem, as with any other problem, sometimes a 'return to basics' is best when all else fails. I posit it starts with philosophy.

Ackermann (Ackermann 1965) paraphrases Plato: '...the senses yield only the appearance of things, but not their reality. To have knowledge about anything we must know it through some other means than the senses.' I would also add that the absence of something either sensible or not can be as important as its expected presence.

Kelly (Kelly 2002) describes Phenomenology-

> 'Phenomenology is essentially descriptive. Its goal is completely and accurately to describe the phenomena of human experience without the interference of metaphysical presuppositions inherited from psychological, scientific, historical, sociological, or other theoretical frameworks.'

Dreyfus (Dreyfus & Dreyfus 2002) introduces a phenomenology of skill acquisition that is divided into five stages with two main groupings. The stages are novice, advanced beginner, competence, proficiency, and expert. The two groups are split at the transition from competency to proficiency. From novice through to competency, these skill levels use 'calculative rationality', that is a reliance upon 'theory and rules', and from proficiency onwards, the use of 'deliberative rationality', that is a reliance upon intuitive experience based skills.

What seems to be expressed here is that rules matter to those who are skilled less than proficient and these theories and rules have less and less relevance as the skill level is raised above the level of competence. From this, experts in a true sense, regard rules not as items that limits actions and responses, but as givens to be considered in the generation of a workable and perhaps a best solution, here, to an IS problem, the real solution relying upon thinking and not rules.

The possible acquisition of IS skills, in my experience, supports, and seems to follow the first three steps of the Dreyfus model's five steps.

- Novices learn rules or non-situational features *–at a training situation and at the start of meaningful employment*;
- Advanced Beginners learn situational aspects through maxims *learned and explained at the workplace*;
- Competence is learned by being fully overwhelmed by theory and rules, and then learning to plan, to cope and by being involved *within the organisation's management and computer systems*;
- The Proficient are making the transition to deliberative rationality where reflecting upon the past and intuition is beginning to replace reasoned rule based responses, (He/she sees what needs to be done, not by rule following alternatives, but by using detached rule following);
- The expert not only sees what needs to be done, but must decide how to achieve the goal. 'The ability to make more subtle and refined discriminations is what distinguishes the expert from the proficient performer'. 'The expert is simply not following any rules!'

(The above from Dreyfus, ibid, the italics are mine)

It must be noted that the 'expert' is not meant to be alone in their decision making processes. To avoid 'tunnel vision' the expert will 'try to protect against this by trying to see the situation in alternate ways, sometimes through reflection and sometimes by consulting others...' even then Dreyfus suggests that 'the expert uses intuition not calculation even in reflection.'

Most importantly, in my opinion, Dreyfus also states 'The increasingly bureaucratic nature of society is heightening the danger that in the future skill and expertise will be lost through over reliance on calculative rationality', the over use of rules and theories to the detriment of thinking about a problem. He concludes 'In general, to preserve expertise we must foster intuition at all levels of decision making, otherwise wisdom will become an endangered species of knowledge.' I posit that this is not being done in IS as well, as the lack of senior non-IS position not being filled by IS personnel shows.

Again, in another paper Dreyfus (Dreyfus 2002a) deals with his five skill levels in this paper as well. Here, he also makes several observations. One is 'To become an expert in any area of expertise one has to be able to respond to the same types of situations as similar as do those who are already expert.' This supports his conclusion that there are two reasons why an expert cannot improve performance by abstract reflection on previous situations, actions, and outcomes.

Firstly, 'that the only way to modify a past response the expert has to feel satisfaction or regret, and not just judge a past action as praiseworthy or condemnable'.

Secondly, 'Detached reflection asks "What was it about the previous action that made it satisfying or regrettable?"'

Here paraphrasing Dreyfus, he adds that since principles were unable to produce expert behaviour for the competent performer, if this performer falls back upon these principles, there would be no surprise if the results were inferior. This is in comparison with the expert who deliberates about the appropriateness of his intuitions. Yet, if the expert is confronted by two or more equally compelling decisions, if the time is available he/she would wait until something is learned that makes one decision intuitively compelling. What is obvious may not be the best decision, and if the problem were presented in a slightly different manner, a different decision may be made. So, the wise decision-maker will attempt to dislodge his current understanding to gain by reflection, a better view of the problem that may yield a better intuitive insight. If this is not possible, then a discussion with other experts may yield the required insight. This means the abandoning of the 'rules' and the acceptance of the non-rationality of intuition, as rules block the development of expertise. In the case of a situation so unlike as any previous situation that no one would have an expert intuitive response, then the expert must go back to detached reflection, that is 'What was it about the previous actions that made it satisfying or regrettable?'

However, the expert is not the final stage of skill acquisition. According to Dreyfus (Dreyfus 2002b), it depends upon the person. He states 'Some people grow up to be experts capable of responding appropriately to a wide rang of interpersonal situations in their culture. Such social experts could be called virtuosi in living.' Of these people, as Dreyfus quotes Pierre Bourdieu, '...others will say, "there was nothing else to be done,"' when examining the problem and its solution at some future date. Also, Dreyfus quotes Heidegger's resolute Dasein, by stating that the individual usually obeys standards and rules, while the resolute individual deviates from the banal, average, public standards to respond spontaneously to the particular situation. The average *responds* to the general situation and the irresolute responds to the concrete by *acting*.

However, all the virtuoso can do is stay open and draw upon his or her experience, and the action resulting can only be *an* appropriate action, not *the* appropriate action. Here we see the difference between the banal that assumes rules protect and the actor who realises that rules have no intrinsic authority and that to accumulate expertise one must take risks.

But there are people who are so radical that they make the marginal practices central and the central (old) practices marginal, these are called Cultural Masters. These are the rare people whose intuition and assessment is their source of industrial influence.

If we accept that the only difference between two installations the same, are the administrative procedures that allow such a system to function, then we are dealing solely with the people who keep the Information System functioning. From this we can make an intuitive conclusion that in a typical installation we have as an essential given, all of Dreyfus' calculative skill levels. What are probably missing are the people with the deliberative skills, the proficient, and the expert and above, not only in any installation, but also in the industry's suppliers as well, who can exercise the necessary power to manage at their relevant skill level

Tobert (Tobert 1991) discusses four time frames of activities in the term that describe the type of power used to manage the activity–

(1) Those that emerge at any time, such as an information attack or 'bug' in the system;

(2) Those that are role defined and last from one week to one year;

(3) Strategic initiatives which last from about 3 years to 5 years and

(4) Mission based activities that last from about 7 years to 21 years, that is from the typical lifetime of a single CEO to a characteristic human generation.

Interestingly, the first two are externally sourced and items 3 and 4 are internally sourced; though all can be, as Torbert put it, interpenetrated. This I take to mean that any combination of the four can affect any or all of the activities of an organisation at any given time.

Summarising, we can evolve two sub-models and three models that describe the situation.

The first sub-model is where calculative rationality exists and where, according to Dreyfus (ibid.), bureaucracy thrives. The problems are externally sourced, short term and where the skill levels of novice, advanced beginner and competent possess the necessary power to resolve general and specialised problems using established procedures. This is where IS typically resides.

The second sub-model is where deliberative rationality exists. The problems are internally sourced, are longer term, and where proficiency and expert skill levels provide the necessary problem resolution skills in the light of the Mission of the organisation.

A first model is where neither the first, nor second sub-models merge and remain as separate entities. I would suggest that this is the usual situation in most organisations today, and in my opinion indicates the beginning of the slow death of the organisation.

A second model is where there is some merging of models either as a permanent or temporary measure. I would suggest that this is where an organisation is trying to evolve into a more effective producer using whatever resources are at its disposal.

A third model is where we combine the first and second sub-model. This is where all problems are resolved using calculative or deliberative rationality or both by an integration of skill levels and all sectors, whether or not that sector or actor is specialised or not. This is supportive of the generation time frame of the current organisational mission and the short term has little or no effect upon the long term unless a change in the Organisational Mission is deemed necessary. This allows IS to participate as an integral part of the whole organisation and not just as the IS section. I consider this the best option of all, but implementing it would be akin to a serious organisational and political evolution.

DISCUSSION

Recently I asked the question 'Are there any experts here?' at a large computer system installation, and the response was 'Technology changes so quickly, we haven't the time to become experts, besides who are the experts in our (the IS) industry?'

If this is true then we have one answer to the riddles of IS security. We cannot out-think our attackers; we can only match them at best, at the competence skill level.

Pressures of costs have limited business to educate to the level of competence in skill acquisition, and then these costs begin again as new technology arrives. The need for experts is shown in the problems of information attack, where the skills needed to save the system from attack are not merely technological knowledge but knowledge about intuitive processes as well.

Though we are dealing with the topic of Information Warfare, the closeness of argument in dealing with the philosophical topics of the Phenomenology of Skill Acquisition matches those of gaining Information Systems skills, up to and including the level of competence. This leads to an idea that for too long IS has been in an encapsulated world without examination about its relationship with the outside philosophical world and how philosophical sources can relate to problems within the IS domain.

There are Information Systems 'experts' who are merely highly competent, having positions requiring experts or better, who are not aware of the their true position. They attempt to resolve problems using, perhaps, inappropriate actions that resolve the problem without consideration of the effects of the 'patch' upon everything else within that organisation. Not only are they perhaps using the wrong tools resulting in a wrong 'patch', they were not trained when to ignore their training when the situation doesn't need rules but needs intelligent thought and reflective practice. They are ignorant of the processes of 'thinking about' and they are offered no help is extricating themselves from this stressful, to themselves and to others, position. So one wonders where do all the highly trained IS people go or have gone after they become comfortably competent?

It seems that one solution is that we make them managers of competency or trainers up to the competency level, as has been done in the past. However, today, we are now faced with new time sensitive problems that mere competency skills may be inadequate to process, so it seems that we are just 'holding the fort', awaiting the next, perhaps ruinous attack.

I posit that the IS industry is in turmoil over security because it does not realise or it is not promoting the idea that there are skill levels above competency. Up to competency, calculative rationality is the rule, the levels above this will be using a more and more sophisticated deliberative rationality the higher the level of skill, perhaps there are commercial reasons for this, or perhaps, the IS industry is ignorant of the philosophy of skill acquisition. I tend to think it is mainly the latter.

A possible solution, perhaps, is special IS training to convert from competency to proficiency to expert. However, it may be that the worst alternative is to do just that. It may be that an IS 'outsider' may be more aware of the realities of being an expert than a highly competent IS technologist. This does not mean that the IS person cannot be trained to expert level. It may be that it would be difficult for that IS person to do so, as they would have to abandon several years of IS training and play 'catch-up' in other than IS organisational training, politics and policies.

CONCLUSION

Competence levels of skills are adequate for the day to day operation and planning of an IS installation, however, the difficulty arises when competency is inadequate for a new problem at hand. Whether it is an Information Attack, a redesign of networks because of evolutionary and revolutionary technology, or a host of other problems, relying upon commercial suppliers is not enough. But where does the user go for the necessary skills, if the industry as a whole is not supportive of the skill acquisition process?

The Dreyfus model of the phenomenology of skill acquisition has been partially explained and how it is applicable to the IS industry. It has been discussed that the IS industry is possibly deliberately

ignorant for commercial or political reasons or does not wish to recognise that there are levels above competency to resolve their IS problems. This does not excuse the user from investing in people or helping people invest in themselves to gain the needed skills to become proficient or even expert in their chosen industry.

Models have been created to show possibilities but realities now come into play. Major rethinking about the role of skill acquisition may not be as difficult as it seems. Many organisations hire 'anybody' and train them in the basics of their business. If they wish to be in IS, they can go there, and if they want to leave IS and do something they consider more effective, they are free to shift. It is that organisation which is split along IS and non-IS lines that will feel the costs of integration of skills and the development of skills up to and including expert. It may be required that an IS person leave his position and join the marketing team before promotion to a higher responsibility level is assured, and perhaps the marketing person learn the intricacies of IS for the same reasons. It is my conclusion that for this to happen in most businesses an evolutionary movement must be started within the IS industry and the industry that employs its skills.

FURTHER RESEARCH

The scope of phenomenology and Information Systems is very broad, and it is in need of more research and training. Perhaps a start is to take one step away from teaching rules to teaching about philosophy.

REFERENCES

Ackermann, R. (1965). *Theories of Knowledge: A Critical Introduction.* New Delhi: Tata McGrawHill Publishing Ltd.

Dreyfus, H. L. (2002a). What is Moral Maturity? A Phenomenological Account on the Development of Ethical Expertise. Available: *http://ist-socrates.berkeley.edu/~hdreyfus/html/papers.html*.

Dreyfus, H. L. (2002b). Could Anything be more Intelligible than Everyday Intelligibility?: Reinterpreting Division I of *Being and Time* in the light of Division II. Available: *http://ist~socrates.berkeley.edu/~hdreyfus/html/papers.html*.

Dreyfus, H. L., & Dreyfus, S. E. (2002). From Socrates to Expert Systems: The limits and Dangers of Calculative Rationality. Available: *http://ist_socrates.berkeley.edu/~hdreyfus/html/papers.html*.

Kelly, S. D. (2002). Grasping at Straws: Motor Intentionality and the Cognitive Science of Skillful Action. Available: *http://www.princeton.edu/~s.kelly/papers/Grasping.pdf*.

Tobert, W. R. (1991). *The Power of Balance: Transforming Self, Society and Scientific Inquiry.* Newbury Park, California: Sage Publications Inc.

The Politics of Cyberconflict

Athina Karatzogianni

University of Nottingham
Nottingham, UK
Email: ldxak1@nottingham.ac.uk

ABSTRACT

The turn of the century witnessed the emergence of a new kind of conflict named 'cyberconflict' to mean conflict in computer-mediated environments (cyberspace). This article seeks to introduce the key terms and themes of cyberconflict and argue that two different trends occur: one between ethnic or religious groups fighting over in cyberspace, as they do in real life and second, between a social movement and its antagonistic institution (hacktivism).

Keywords: Internet politics, cyberconflict, hacktivism

INTRODUCTION

The term 'cyberconflict' is now in regular use, but has not been sufficiently clarified as yet. This is because there are problems in defining/categorising the wide variety of events occurring in cyberspace that fall under this conceptual umbrella. The political use of the Internet has created a new lexicon. The term cyberwar is used to refer to the destruction of the enemy's infrastructure through information technology and the term netwar refers to conflict between network-type groups using information technology to organise and promote their political agendas. Here, the term cyberconflict is used to refer to conflicts of the real world spilling over to cyberspace, in which the opposing parties use either Information Technology as such or IT as a weapon-for example worms, Distributed Denial of service attacks (DDos), Domain Name Service attacks (DNS), or unauthorised intrusions -to attack the other side.

Most contemporary viruses are worms, enabled by 'buffer overflows'. Buffer overflow is an event in which more data is put into a buffer (computer holding area) than the buffer has been allocated. In recent years they have become very popular with new worms such as Code Red, Code Red II, which allows it to gain control of the infected machine and the Nimda worm. (Vatis, 2001). Some researchers have predicted the emergence of new worms (Warhol worms, flash worms), spreading in seconds, leaving no time for system administrators to react. (Weaver, 2001). DDos attacks employ armies of 'zombie' machines (insecure server compromised by a hacker who places software on it that can launch an overwhelming number of requests, rendering the site inoperable). As demonstated in the Kosovo conflict, military web sites and communications systems are especially likely to receive DDos variants (Vatis,2001). Domain name servers are the 'Yellow Pages' that computers consult in order to obtain the mapping between the name of a system and the numerical adress of the system. If the DNS provides an incorrect numerical adress for a website then the user's system will connect to the incorrect server. An attacker can disseminate false information this way and prevent acces to the original site (Vatis,2001).

These type of attacks are called cyberattacks. The argument of this paper is that cyberconflict includes two different categories, which are sometimes blurred. The first is between two ethnic or religious

groups that fight it over in cyberspace, as they do in real life and the second is between a social movement and its antagonistic institution.

It is not surprising that the Internet has been used vigorously by social activists and campaigners all over the world. The Internet quickly puts information into the hands of organisers, allows rapid replication of a successful effort, allows users to select their level of activity and helps publicise the campaign. It is therefore an organising tool *par excellence*, because the more traditional telephone trees or fax machines are two slow and the physical distances are too difficult and too expensive to cover (Danitz and Strobel, 2001, p.162). However, the Internet is not only used by social activists. An examination of historical precedents indicates that major political and military conflicts are increasingly accompanied by a significant amount of aggressive activity. Ongoing conflicts also show that cyberattacks are escalating in volume, sophistication and coordination (Vatis, 2001). Parties in cyberconflicts have been described as terrorists or social activists depending on the discursive mood of their critics. This is why it is important to examine the politics of this phenomenon and understand its implications for future conflicts.

Netwarriors: Terrorists or Social Activists?

Arquilla and Ronfeldt (2001) differentiate between the terms 'cyberwar' and 'netwar'. While cyberwar refers to a more 'heavy' mode of new military conflict like destruction of the enemy's infrastructure through information technology, the term netwar was devised to refer to information age conflict at the less military, low intensity, more social end of the spectrum. According to these writers, there are roughly two categories of netwarriors: those who are violent and negative, i.e terrorists and criminals, or social activists, who maybe militant but can be also peaceable and even promising. The definition of netwar Arquilla and Ronfeldt use is the following:

The term netwar refers to an emerging mode of conflict (and crime) at societal levels, short of traditional military warfare, in which the protagonists use network forms of organisation and related doctrines, strategies and technologies attuned to the information age (Arquilla and Ronfeldt, 2001, p.6).

One of the interesting sides to their argument is that the information revolution is altering the nature of conflict by strengthening network forms of organisation over hierarchical forms and that the conduct and outcome of conflicts increasingly depends upon information and communications. Following that argument, Arquilla and Ronfeldt go on to say that hierarchies have a difficult time fighting networks (e.g. Colombia, Algeria, Zapatistas). It takes networks to fight networks and whoever masters the network form first and with the most success will gain major advantages (Arquilla and Ronfeldt, p.55).

Terrorists, when operating with an organisational networking system, are likely to be using IT to help co-ordinate and support their activities, not solely as a weapon in the case of cyberwars. There is no need to repeat here why IT is the next best thing (quicker, cheaper, secure, anonymous and would be covered by the media because of the public's fascination by computer attacks), but it would be useful to stress that terrorists might be more interested to keep the Internet up and running than to disrupt or destroy its components. Moreover, one has to agree with Arquilla and Ronfeldt that networked based conflict will become a major phenomenon of the future and this can be clearly demonstrated with Al–Qaeda even long before the September 11th attack with reports of Bin Laden having advanced information systems designed by Egyptian computer scientists. As Arquilla and Ronfeldt (2001) argue:

These protagonists are likely to consist of dispersed small groups who communicate, co-ordinate and conduct their campaigns in an internetted manner, without a precise central command....To give a string of examples, netwar is about the Middle East' s Hamas more than the Palestine Liberation Organisation (PLO), Mexico's Zapatistas more than Cuba's Fidelistas, and the American Christian

patriot movement more than the Ku Klux Klan. It is also about the Asian triads more than the Sicilian Mafia, and the Chicago's Gangsta disciples more than the Al Capone Gang (p.45).

Hamas in the United States uses Internet chat rooms and e-mails to co-ordinate their activities across Gaza, the West Bank, and Lebanon, making it difficult for Israeli security officials to trace their messages and decode their contents (Denning, 2001).
In a 1998 US News and World Report there was evidence of 12 of the 30 groups on the US State Department's list of terrorist organisations on the Web. More recently, it seems that virtually every terrorist organisation is on the Internet (Denning, 2001, p.252).

Furthermore, launching a cyberattack is fairly inexpensive. One knowledgeable hacker with a computer can wreak havoc on an automated pipeline. A cell of 'cyber-space guerrillas' armed with a few thousand dollars' worth of hardware could disable a nation's power grid. Several hackers together can dramatically increase the capabilities of a terrorist group (Jane's international police issues, 1998). More guerrilla groups will be attracted to cyberwarfare, because they can spread propaganda, recruit sympathizers and collect data. The possibilities are endless: military espionage, control and disruption of information flow, destruction, distortion and fabrication of data, electronic bombs and psychological operations could be potential tactics.

A report released on May 1999 from the Rand Corporation concluded that the Airforce of the US should slow its modernization plans and rethink its connections to the Internet if it wants to fight off a revolutionary, yet undeveloped form of cyberterrorism. According to the report, which was the result of a yearlong project sponsored by the Airforce's Deputy Chief of Staff for Air and Space operations, there will be a new form of terrorism known as 'netwar':

The rise of networks is likely to reshape terrorism in the Information Age and lead to the adoption of netwar- a kind of Information Age conflict that will be waged principally by nonstate actors. There is a new generation of radicals and activists who are just beginning to create Information Age ideologies. New kinds of actors, such as anarchistic leagues of computer- hacking 'cyboteurs' may also partake of netwar (Verton, 1999).

The report also predicts that cyberterrorists will use new tactics such as 'swarming', which occurs when members of a terrorist group, spread over great distances, electronically converge on a target from multiple directions, a tactic different from the traditional form of attacking in waves, which delivers a knock out blow from a single direction on the Internet (Arquilla and Ronfeldt, 2000).

Winn Schwartau, a computer security expert, argues against those who believe that information warfare is just a hype, that cyberattacks can cause terror. An example he gives is waging an Information Warfare attack against a series of US domestic air carriers where you do not use a conventional bomb but an electromagnetic bomb, which sufficiently interferes with the avionics of the plane to cause it to crash (Venke, 1996). As early as 1995 the Pentagon's assistant director for strategic planning made clear that the problem is real enough:

...as the information age matures, a truly revolutionary form of warfare will emerge. Information warfare will be fought in a different environment, with adversaries grappling in cyberspace. As every potential adversary achieves access to multiple information systems, warfare will be conducted virtually at the speed of light over global distances. Domination of cyberspace may render the need to employ conventional forces and firepower less likely (Lt. Tod as quoted in Guisnel, 1997, p.178).

CYBERCONFLICT: The Socio-political Dimension

Netwar conflicts can be placed into two broad groups, either those who are primarily concerned with global issues such as the environment, issues for which the level of negotiation with governments is open for debate, and those groups who are much less inclined to negotiate with governments, being as

they are concerned with issues such as their own liberation from the control of the state (Arquilla and Ronfeldt, 2001, p.15). The main purpose of these activists is to influence or challenge public opinion, or the enemy and battle for media access and coverage.

There is a debate between hacktivists concerning denial of service attacks and web defacements. On the one side there are those who find that such actions run contrary to other people's right to freedom of speech and those who view these actions as the only way to get the public' s attention. The fact is that Web defacements cannot be dismissed as electronic graffiti and denial of service attacks as nuisances, because there is concern by online companies that it could affect a significant share prices, earnings and cause damage to reputation and customer confidence. Indeed this is the reason why there should be more analysis on the reasoning underpinning hacktivism and its political rationale.

One aspect of the problem seems to be that with the explosion of the size of the Internet, protests and political activism have entered a new realm. Political activism on the Internet has generated a wide range of activity such as using e-mail and web sites to organize, to web defacements and denial-of-service attacks, described above. These politically motivated attacks are defined as 'hacktivism'. Stanton McCandlish, program director of the Electronic Frontier Foundation describes it as follows:

A kind of electronic civil disobedience in which activists take direct action by breaking into or protesting with government or corporate computer systems. It's a kind of low-level information warfare, and it's on the rise (Phrack, 1998).

Tim Jordan writes that social activists or hacktivists have found two uses for the Internet: Mass Virtual Direct Action (MVDA) and Individual Virtual Direct Action (IVDA) (Jordan, 1001, p.8). According to Jordan, MVDA involves the simultaneous use by many people of the Internet to create electronic civil disobedience. There are two characteristics of this type of hacktivism: First, hacktions are not aimed at halting a target permanently, but have symbolic dimensions. Second, MVDA activists rarely try to hide their identities, seeking public debate and discussion. Individual Direct Virtual Action differentiates from MVDA in that it could be taken by an individual and does not depend on a mass protest. The actions taken are either semiotic attacks (i.e. defacements), computer intrusion or network security.

An example of hacktivism would be the Seattle anti-WTO protests at the end of November 1999, which were the first to take full advantage of the alternative media network via the Internet. Protestors used cellphones, direct transmissions from independent media feeding directly onto the Internet, personal computers with wireless modems broadcasting live video, and a variety of other network communications. This is how Paul De Armond (2001) describes it:

Floating above the tear gas was a pulsing infosphere of enormous band width reaching around the planet via the Internet - although on the scene, at street level, the Internet played little role, because most communications among the affinity groups were face to face and via cell phone, unencrypted (p.210).

During the Seattle anti-WTO protests, hackivists managed to acquire the URL www.gatt.org, using the GATT address for a parody WTO site, looking identical to the original WTO one, but included a text criticizing WTO trade policies. Another famous hacktion was a MVDA by the Electrohippies, which included a virtual sit-in with a downloadable web page aiming to flood the WTO's server. The Electrohippies claim that 400,000 hits in this MVDA had slowed down the WTO and at times completely halted it (Jordan, 2001, p.8).

Another example of Mass Virtual Direct Action on the Internet is the one organised by the Electronic Disturbance Theatre (EDT). To demonstrate solidarity with the Zapatistas, an estimated 10, 000 people from all over the world participated in the sit-in on September 9, 1998 against the sides of President Zedillo, the Pentagon and the Frankfurt stock exchange, delivering 600,000 hits per minute to each. The EDT have explained their action accordingly:

We do not believe that only nation states have the legitimate authority to engage in war and aggression. And we see cyberspace as a means for non-state political actors to enter present and future arenas of conflict and to do so across international borders (Denning, 2001, p.267).

The Internet was crucially influential in enabling civil society actors to force the passage of a series of laws regarding business and political dealing of the US with Burma, resulting in a Massachusetts decision forbidding companies to do business with Burma and a reaction from Europe and Japan against these laws. The result was that relatively 'insignificant' constituents in the US were able to influence American foreign policy using the Internet, a form of what might be called cyberdemocracy. As Danitz and Strobel describe it, both sides in the conflict, the SLORC (The Burmese government) and prodemocracy advocates have long been engaging in an information conflict in an attempt to influence international public opinion. The prodemocracy activists have used the Internet in the Massachusetts campaign, US citywide selective purchasing campaigns and the Boycott Pepsi campaign and also to inform journalists of forthcoming events, Burmese problems like slave labour and government oppression. SLORC has produced a web page but has not taken full advantage of the technology (Danitz and Strobel, 2001, p.142).

Human rights workers increasingly use the Internet to co-ordinate their actions against repressive governments. Encryption is one of the tools that they use because it allows them to protect communication and stored information from government interception. For example, human rights activists in Guatemala attributed the saving of lives of witnesses to military abuses to the Pretty Good Privacy tool (PGP) (Denning, 2001: 258).

CYBERCONFLICT: The ethno-religious dimension.

The increasing importance of cyberconflict is even more evident when it reflects conflicts belonging to the real world. In October 2000, Israeli and Palestinian hackers engaged in adversial hacking when the prolonged peace talks between the two parties broke down. Until the beginning of November 2000 groups supporting either side in the conflict limited their online activities to defacements and denial-of-service attacks against websites affiliated with the Palestinian movement or Israeli nationalists. One example was when an Israeli flag, Hebrew text and a piano recording of 'Hatikva', the Israeli national anthem, appeared on the Hezbollah home page (Hockstader, 2000).

Also, Palestinian hackers created a web site called Wizel.com - a host for FloodNet attack, which reloads a targeted web page several times, which makes the site inoperable. The reaction was a sustained counterattack from Pro-Palestinian 'cybersoldiers' from the US. The web sites of the Israeli Army, Foreign ministry and Parliament among others were attacked. Targets included financial institutions; e-commerce sites crashed and there was an economic impact reflected in the Israeli Markets.

The situation however escalated in the first days of November 2000, when an anti-Israeli hacker attacked the website of one of Washington's most powerful lobbying organisations, the American-Israeli Public Affairs Committee (Aipac). The hackers published critical e-mails downloaded from Aipac's own databases and credit card numbers and e-mail addresses of Aipac members. After the FBI was informed the members of the organisation, including a Republican Senator were advised to cancel their credit cards and monitor their accounts. The hackers wrote: 'The hack is to protest against the atrocities in Palestine by the barbaric Israeli soldiers and their constant support by the US government' (BBConline, 3/11/00). Aipac spokesman Kenneth Bricker at the time said that the hackers downloaded credit card numbers and about 3,500 names and web addresses from people who had contracted Aipac's web site. The broadest list of the organisation's 55,000 members were stored in a separate computer system and were not compromised.

The Israelis were not slow to retaliate. According to MAGLAN, an Israeli information warfare research lab, an Israeli supporter, 'Polo0', posted Palestinian leaders' cell phones numbers, as well as

information about accessing the telephone and fax systems of the Palestinian Authority, plus 24 different web sites, 15 IRC channels and an IRC server through which the Palestinian movement communicates. Analysis by iDefense, a security monitoring agency, considered a number of key players in the cyber conflict. On the Israeli side the wizel.com creators, a.israforce.com, Smallmistake and Hizballa attacked Palestinian sites. On the Palestinian side there is Unity, a Muslim extremist group, one of the forerunners as to what is referred to as 'e-jihad' or 'cyberjihad'. Unity attacked the Tel Aviv Stock Exchange. Later they announced that their strategy was four phased. Phase one included crashing official Israeli government sites, phase two hit the Bank of Israel and phase three targeted the Israel ISP infrastructure, Lucent and Golden airlines, an Israeli telecommunications provider. They also said that they would not realise phase four, the destruction of e-commerce sites, but added: 'We warn the Zionists and their supporters that any attempt to touch any Anti-Zionist site will be faced with phase four of the cyberwar-causing millions of dollars in transactions' (Gentile, 2000). Unity also claimed in an e-mail in February 2001 to have successfully attacked AT&T in retaliation for the company doing business-as a back up in case of emergency- with the Israeli Defence Force, claiming to have blocked the site for seventy two hours in a particular hit. AT&T declined to comment at the time. (Galvin, 2001)

Fred Cohen, a computer-security professor commented at the time: 'When you talk about war, you are talking about turning off the constraints that hold back people. You have people who want to break into computers, and now they have an excuse -they can do it for a cause' (Lemos, 2000).

What distinguishes this cyberconflict from past ones is that it moved beyond being a game of high specialised hackers into involving thousands of Israeli and Arab youngsters sending racist and occasionally pornographic e-mails and within their own camps circulating Website addresses with simple instructions for how to crush the enemy's electronic fortresses (Hockstader, 2000). One site offered a menu of targets to attack, including the sites of Hezbollah, the Palestinian national Authority, Hamas and a dozen others. The site said: 'Come and help us stop their pan-Arabic campaign of incitement. Our purpose is not to allow the cruel terror organisations to continue with their spreading terror, articles and sick pictures throughout the Internet' (Hockstader, 2000). The site then encouraged users to click on the targets they would like to disable and offered a set of simple instructions for executing the assault. The whole process did not take more than a minute or two and generated multiple and high speed attacks.

IDefense's director of intelligence production Ben Venzke thinks that the Palestinians won this particular battle in cyberspace, because according to him, there are people on the Palestinian side trying to learn how to hack overnight to join the effort. If quantity is a measure of success the Palestinians seem to be winning the Inter (Net)-Fada: They had struck over 166 web sites during the months of October-January 2001, while at the same time the Israelis had hit approximately 34, according to an iDefense report, which came out on June 3rd 2001 (Hershman, 2001). According to the report the cyberconflict would intensify as political tensions in the region heighten. Venzke referred to a hacker called Dodi, responsible for some of the most destructive attacks in this war. On November 3rd 2000, Dodi after defacing an Israeli service provider said he could shut down NetVision and added: '...this is not just a war against Israel, for the perpetrators of the atrocities in Palestine are US-backed. It's America which has blood on its hands, the blood of innocent women and children' (Gentile, 2000).

During a conference in Munich in June 2001 titled: Cyberwar between Israel and Palestine, Dan Caspi of the Ben Gurion University said that his research on the Al Bawba website and its chat room gave him a sour view:

When you enter into the chat, you immediately feel very attacked, as a person and as a representative of your society. So you don't have any choice but to adopt a kind of role-playing you normally don't agree with. So this forum normally contributes to the polarisation of debate. When I started, I was very optimistic to see one site where ordinary Jews could talk to ordinary Arabs. But after a few days I can tell you there is only a small community and they don't allow you to join them (Kettman 2001).

The political crisis in the beginning of 2002 in the Middle East spawned again an increase in defacement attacks on Israeli Web servers. Israel was the victim of 10 out of 15 significant web defacements in the Middle East over the first two weeks of April 2002 according to security consultancy mi2g . Mi2g reported Israeli web sites with the "il" domain were defaced 413 times in 2001- up 220 per cent from the year before and has been the biggest victim of web defacements over the past three years, suffering 548 of the 1,295 attacks in the Middle East. The most active anti-Israeli hacker group claims to be Egyptian and started its activities just after September the 11th (Leyden, www.mi2g.com).

Also, sympathizers on both sides of the Kashmir conflict have used cyberattacks to disrupt each other's computer systems and disseminate propaganda One of the first moments of cyberwarfare in the region was reported on the16th of October 1998 by the Indian news agency PTI. They said, 'suspected Pakistani intelligence operatives had hijacked the Indian Army's only website, Kashmir A Paradise, which gives the Indian view on Kashmir' (BBConline, 25/10/1998). The site was set up a month earlier as counter-propaganda to the dozens of sites supporting Muslim Kashmiris seeking independence. Among the propaganda there was a guest book where visitors can leave comments. Two typical responses from the opposing sides were: (Pakistani) 'This web site is very biased and very unfair to the Pakistani point of view. This is just a whole charade by the Indians and 80% of it is absolutely untrue!' (Indian) 'A whole hearted salute for my brothers fighting for our country with a religious maverick enemy' (Nuttal, 1998). The hackers had put information on alleged torture of Kashmiris by the Indian security forces.

The attack came at the time when Indian and Pakistan began talks in Islamabad in an effort to ease tensions. The Pakistani hackers dedicated the 'new' site to "all the Kashmiri brothers who are suffering the brutal oppression of the Indian army" (BBC online, 16/10/98). The photographs of the site were overwritten with the slogans: 'Stop the Indians' and 'Save Kashmir'. Pictures showing Kashmiris allegedly killed by Indian forces were posed under headings such as 'massacre', 'torture', 'extra-judicial execution' and 'the agony of crackdown'. A government statement said the hackers changed the site parameters so that visitors were diverted to a different server. In March 2000, the cyberconflict escalated when a group of Pakistani hackers defaced 600 websites and took over temporarily government and private computer systems. The majority of the sites were hacked after the Pakistanis broke into IndiaLinks, India's largest Internet service provider.

The team responsible were the 'Muslim on Line Syndicate', described by their spokesperson as a group of nine ranging from 16 to 24 years of age. Their spokesman also described their operations as taking control of a server, then deface the site, after they have no more use for the data or the server. Their message was: 'We hope to bring the Kashmir conflict to the world's attention... We wish that our Muslim brothers will be given the right to choose, as was promised them half a century ago' (Hopper, 2000). The number of Pro-Pakistani defacements of Indian web sites has increased dramatically between 1999 and 2001: 45 in 1999, 133 in 2000 and 275 by the end of August 2001(The Statesman, 2001). However, the assault on Pakistani sites has not been as successful. There were reports that they have repeatedly tried to hack a Pakistani newspaper called Dawn, without any result. Nevertheless they have left messages to their Pakistani counterparts like 'keep your hands off Indian sites', threats of 'breaking the Internet backbone' of Pakistan and that 'India is the superpower of Information Technology' (Joseph, 23/12/00).

Another example is the cyberconflict emerging during an international diplomatic incident. When a US spy plane made an emergency landing on Chinese soil on April 1st 2001, after colliding with a Chinese fighter jet over the South China Sea, killing the Chinese pilot, Chinese hackers vowed to attack US sites, which led hackers in the US to retaliate.

According to UK computer security firm Mi2g, the Honkers Union of China hacking groups defaced 80 websites and the Americans defaced more than 100 during April 2001 (Left, 2001). China's remote sensing satellite ground station was overwritten with a picture of a mushroom cloud, while in the US, the White House historical association was plastered with Chinese flags as were the departments of

Health, Navy, Labour, as well as the House of representatives' e-mail servers. On May 9th 2001, Chinese hackers boasted they had defaced 1,000 US web sites, but called a truce to the conflict. A statement by the Honker Union of China said that having attacked 1,000 sites they reached their goal and that any attack from that point on had no connection with them. Their American counterparts broke into hundreds of Chinese sites, leaving messages such as: 'We will hate China forever and will hack its sites' (Globe Technology, 10/5/01). After a meeting online between Honker union and the Chinese Red Guest Network Alliance, where it was decided that their attack would last a week, ending on May 7th, the two-year anniversary of the bombing of the Chinese embassy, they decided to keep the destruction of business websites to a minimum and attack instead government web sites. They said that the point of the attack was to encourage people in the US to protest against their government and demand peace between nations. One hacker said:

The U.S. wants the world to go to war. All people cherish peace, but the mildew dog government of the U.S wants war. We will attack to send a message to the people of the US, to tell them we are all one, but they must stop their government from destroying the world. (Delio, 2001).

Attacks that were discussed on an Internet Relay Chat during their meeting involved defacing web sites, e-mailing viruses to U.S government employees and flooding computers with garbage data. A US hackers collective dubbed Project China left this message on a Chinese site: 'Get ready to meet a strike force with strength the world has never seen before! We are going for all-out cyberwarfare on your gov.cn boxed and every box that you fucks haven't secured!' (Left, 2001). The Xinhua News Agency reported at the time that U.S hackers had defaced the web sites of the provincial governments of Yichun, Xiajun and Beijing, the Deng Xiaoping Universities, and Samsung's and Daewoo Telecom's Korean sites. A South Korean government security agency blamed the Sino-US cyberwar for the 164 cyberattacks on South Korean Websites that had occurred during that time. Computer analysts said that American and Chines hackers were using Korea to get into rival countries' computer systems without revealing their identities, because S.Korea has extensive links with both countries.

Interestingly, the Chinese government has been quite open about its future strategic military objective. In the 2001 spring issue of the China Military Science Journal, a member of the Chinese Committee of Science, Technology and Industry of the System Engineering Institute wrote:

We are in the midst of a new technology in which electronic information technology is the control technology. The technology provides unprecedented applications for the development of new weaponry...Military battles during the 21st century will unfold around the use of information for military and political goals (Chepsiuk, 2001).

CONCLUSION

Cyberconflict should not be dismissed as just catfights between computer geeks. Mi2g Chief executive, DK Matai has argued that cyberwarfare could be used as a barometer for political tensions around the world: 'The tense situation in the Middle East is reflected in both covert and overt hack attacks' (BBConline, 16/4/02). In addition, cyberconflict is a phenomenon that includes a variety of actors with different characteristics, which cannot be easily distinguished as either terrorist or activist in nature. Accordingly, the political game between parties in an ethno-religious conflict or among social activists engaged in a socio-political one is neither clear, nor fully developed. Nevertheless, it has to be stressed that we are discussing two different kinds of cyberconflict and they should be treated as such. A suggestion would be to use conflict theory to analyse ethno-religious type cyberconflicts and new social movement theory to examine social activism on the Net.

REFERENCES:

Arquilla, J and Ronfeldt, D: (2000): *Swarming and the Future of Conflict*, California: Rand.

Arquilla, J and Ronfeldt, D (eds.): (2001) *Networks and Netwars: The Future of Terror, Crime and Militancy*, California: RandBBConline: War of words on the Internet, 25/10/98

BBConline: Indian army web site ambushed, 16/10/98

BBC online: Israeli lobby group hacked, 3/11/00.

BBC online: Israel under hack attack, 16/4/02.

Chepsiuk, R: Get ready for Cyberwars, *ncmonline*, 23/8/01

Danitz, T and Strobel, W: Cyberactivists use the Internet to promote democracy in Burma, in Arquilla and Ronfeldt (2001) *Networks and Netwars: The Future of Terror, Crime and Militancy*, California: Rand.

Delio, M: It's Cyberwar: China vs U.S., Wired News, 30/4/01

Denning, D: Activism, Hacktivism and Cyberterrorism: the Internet as a tool for influencing foreign policy, in Arquilla and Ronfeldt (2001) *Networks and Netwars: The Future of Terror*, Crime and Militancy, California: Rand.

De Armond, P: Netwar in the emerald city: WTO protest strategy and tactics, in Arquilla and Ronfeldt (2001) *Networks and Netwars: The Future of Terror, Crime and Militancy*, California: Rand.

Galvin, J: 'Cyberwars bring real world conflict to the Web', *ZDNet,* 20/2/01.

Gentile, C: 'Hacker War Rages in Holy Land', *Wired News*, 8/11/00

Globe Technology: Chinese hackers halt Web war, say 1,000 sites defaced, 10/5/01

Guisnel, J: (1997) *Cyberwars*, New York and London: Plenum Trade.

Hershman, T: Israel discusses 'Inter-Fada', *Wired News*, 29/6/01

Hockstader, L: Pings and E-Arrows fly in Mideast Cyber-War, *Washington Post online*, 27/10/00

Hopper, I: Kasmir conflict continues to escalate online, *CNN.com*, 20/3/00.

Jane's International Police issues: Information warfare: the coming threat, 2/12/98

Jordan, J: Mapping Hacktivism, *Computer Fraud and Security*, 2001

Joseph, M: Both sides hacked over Kashmir, *Wired News*, 23/12/00.

Kettman, S: Deep thinking on the 'Inter-Fada', *Wired News*, 10/1/01

Lemos, R: Hacktivism: Mideast Cyberwar heats up, *ZDNet news*, 5/11/00

Left, S: Chinese and American hackers declare cyberwar, *Guardian unlimited,* 4/5/01.

Leyden, J: Middle East conflict spills over into cyberspace, *the Register, www.mi2g.com.*

Nuttal, V: India opens virtual front in Kashmir, *BBC online* 5/10/98.

Phrack Magazine, vol.18, no 54, Dec 1998

The Statesman: Pro-Pakistan hackers deface centre's venture capital site, 21/8/01

Vatis, M: (2001): *Cyberattacks During the War on Terrorism: A Predictive analysis*, Dartmouth: Institute for Security Technology Studies, Dartmouth College.

Venke, B: Information Warrior, *Wired*, 4/8/1996

Verton, D: New Cyberterror threatens AF, *Federal Computer Week*, 3/5/99

Weaver, N: *Warhol Worms: The potential for very fast Internet Plagues*, California: University of California Berkley, Aug.15, 2001

Biometric Authentication for Mobile Devices

N.L. Clarke[1], S.M. Furnell[1] & P.L. Reynolds[2]

[1]*Network Research Group*
Department of Communication & Electronic Engineering
University of Plymouth, Plymouth
United Kingdom
Email: info@network-research-group.com

[2]*Orange Personal Communications Services Ltd*
Bradley Stoke, Bristol
United Kingdom

ABSTRACT

Mobile devices have found an important place in modern society, with hundreds of millions currently in use. The majority of these use inherently weak authentication mechanisms, based upon passwords and PINs, which can potentially be compromised and thereby allow attackers access to the device and its stored data. A need for stronger authentication is identified and the discussion considers the application of various biometrics to a mobile platform. The feasibility of one such approach, that of keystroke dynamics, is examined, revealing promising results – with individual performances of 0% false rejection rate and 1.3% false acceptance rate being observed. However, higher overall error rates of 15% lead to the proposal of a hybrid, non-intrusive approach to authentication.

Keywords: Mobile Devices, Authentication, Biometrics

INTRODUCTION

The ability to communicate and work whilst on the move has given rise to an explosive growth in mobile devices. Primarily this growth has come out of mobile phone related technologies with worldwide subscribers now in excess of a billion (UMTS Forum, 2002), but it can also be seen that both the use of Personal Desktop Assistants (PDA's), and laptop computers has been growing with popularity (Richardson, 2002; Gibson, 2001). However, this rise in computing mobility could cause a number of security issues, in particular with attackers accessing the data stored on the devices.

The most popular access security to date takes the form of the password or PIN (Personal Identification Number), a secret-knowledge approach that relies heavily on the user to ensure continued validity. For example, the user should not use the default factory settings, tell other people, or write it down. However the poor use of passwords and PINs has been widely documented, with many laptops owners using simple passwords that dictionary attacks can crack in seconds and with many mobile phones and PDA users not even using the security available. Recent surveys have indicated that 44% of mobile phone users do not use the PIN and 25% of PDA users do not a password (Clarke et al., 2002a; Leyden, 2002). Taking a crude comparison with current mobile phone subscribers, this would indicate that some 500 million mobile phones have no access security. Although this is not a particular issue currently with the second generation mobile phones with their

limited storage and computing abilities, this will change with the advent of third generation networks and a convergence of PDA and mobile phone functionality (Giussani, 2001). Mobile phones will be able to store detailed information about friends and family, include digital certificates, bank details and be able to access a wide range of data services through your phone account – ranging from the purchasing of goods to watching movies. Interestingly the same mobile phone survey found that, in contradiction to not using the protection already available with 41% of respondents citing inconvenience, that 81% of respondents wanted more security.

So an alternative means of subscriber authentication is required to replace the secret-knowledge based approaches. It is therefore appropriate to examine the potential of a fundamentally different strategy. From the available techniques, that of token-based authentication and biometric based authentication, only the latter really seems plausible, since tokens would also have to be carried with you along with the device or more commonly left permanently in situ. Biometrics, are based not on what the user *knows*, or what they carry, but who the user *is*, some unique characteristic. After explaining the biometric concept in more detail, this paper considers the techniques that could potentially be deployed on mobile devices, along with a brief example of a practical implementation.

THE NEED FOR AUTHENTICATION ON MOBILE DEVICES

As previously indicated, a large number of mobile devices are currently in use with little or no authentication security. A recent survey into the use of PDAs discovered a third of users who have already had their PDA stolen once still do not use a password, however, one of the cited uses for a PDA by respondents was to store all the passwords and PINs they regularly use for other systems (Leyden, 2002). This highlights two primary issues; firstly, the inherent weaknesses of secret-knowledge based techniques such as the password in that they can be written down in the first place, and secondly the importance of the data being stored on the device. There is a third issue raised concerning user perception and realisation of the security problems. Any person storing sensitive information on a device without securing that device clearly has little comprehension of the associated security issues.

The security weaknesses and threats associated with PDAs are important because although the number of devices currently in use is relatively small (in the order of tens of millions), the mobile phone is set to absorb and surpass much of the functionality of current PDA devices. The difference in numbers is from tens of millions of PDAs to hundreds of millions of mobile phones. If authentication mechanisms were left as they currently stand, then the threat posed by attackers would inconvenience users through cost associated with misuse and an almost certain increase in the theft of the devices. For example, the UK Home Office reported some 700,000 mobile phone thefts from subscribers in 2001 and this number can only be set to increase as mobile phones are packed with more technological wizardry (Harrington et al, 2001).

Concerns can also be expressed in relation to laptop computers. For example, the UK Ministry of Defence (MoD) admitted to losing over 600 laptops over a five year period (BBC, 2002), many obviously containing very sensitive information. Although it is likely that many laptops are stolen merely to be resold as a piece of equipment, rather than for the information stored upon them, this cannot be the case it all thefts. Infosecurity reported in May 1999, that 57% of computer crimes involving break-ins on corporate servers were linked to stolen laptops that enabled the breach (Broomfield, 2000).

BIOMETRIC APPROACHES & IMPLEMENTATION

The use of biometrics has existed for hundreds of years in one form or another, whether it is a physical description of a person or perhaps more recently a photograph. Consider for a moment what it is that

actually allows you to recognise a friend in the street or allows you to recognise a family member over the phone. Typically this would be their face and voice respectively, both of which are biometrics. Biometrics are based on unique characteristics of a person, and are typically subdivided into two categories, physiological and behavioural. Physiological biometrics are those based on classifying the person according to some physical attribute, such as their fingerprints, their face and their hand. Behavioural biometrics rely on a unique behaviour of the person such as, their voice and the way in which they write their signature.

Biometrics all work on the basis of comparing the biometric sample against a known template, which is securely acquisitioned from the user when he or she enrolled on the system initially. However this template matching process gives rise to a characteristic performance plot between the two main error rates governing biometrics. The False Acceptance Rate (FAR), or rate at which an impostor is accepted by the system, and the False Rejection Rate (FRR), or rate at which the authorised user is rejected from the system. The error rates share a mutually exclusive relationship as one error rate decreases, the other tends to increase, giving rise to a situation where neither of the error rates are typically both at zero percent (Cope, 1990). Figure 1 illustrates an example of this relationship.

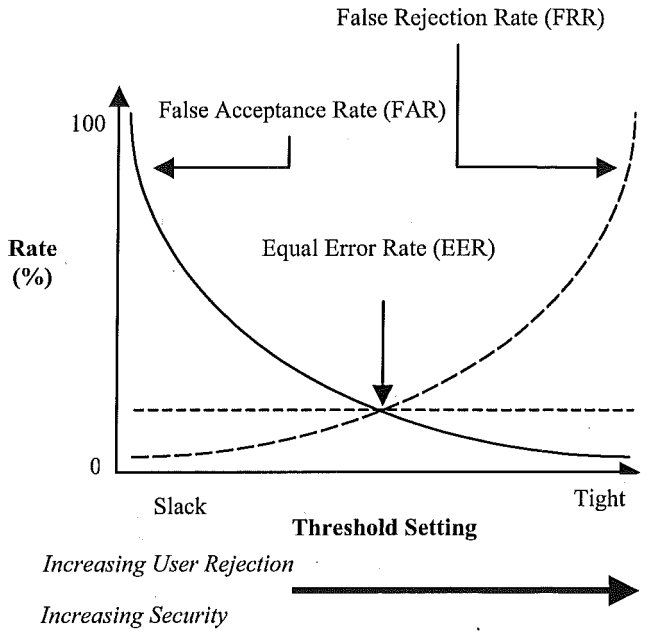


Figure 1 Mutually exclusive relationship between the False Acceptance & False Rejection Rates

This leads to a trade-off situation between high security and low user acceptance (due to fact the authorised user is being rejected a large proportion of the time) and low security and high user acceptance, to which a decision has to made about what threshold setting to set that meets both the security requirements of the device and acceptance levels of users. The point at which the error rates cross is called the Equal Error Rate and is used in industry as a comparative measure between different biometric approaches (Ashbourn, 2000).

The next section provides an overview to the most common biometrics that could be implemented within a mobile terminal, indicating what the unique characteristic the technique attempts to classify users upon and how the biometric is obtained. For more general information on any of the approaches discussed here, consult Nanavati et al. (2002) and Smith (2001).

PHYSIOLOGICAL BIOMETRIC

- Fingerprint Recognition

The most commonly deployed biometric, with a mature and proven technology. The fingerprint comprises of ridges and valleys that form distinctive patterns, such as loops, swirls and arches. The ridges and valleys are characterised by discontinuous and irregularities known as *minutiae* – these are the distinctive features on which most fingerprint technologies are based. In order for the fingerprint image to be captured a specialist reader is required

- Facial Recognition

This utilises the distinctive features of the human face in order to authenticate a user. The features often used are those which change very little over time, such as the upper ridges of the eye sockets, areas around the cheekbones, sides of the mouth, nose shape and the relative position of these features relative to each other. The facial image itself can be generated from any static camera or video system that is able to generate image of sufficient quality, such as web camera.

- Iris Scanning

Iris scan technology works by utilising the distinctive features of the human iris and has the potential to be one of the most successful biometrics (Harrison, 2001). Iris recognition requires the acquisition of a high-resolution image of the eye, illuminated by an infrared imager, in order to effectively map the details of the iris. The device to capture this image can vary from a desktop camera to a dedicated camera for integration into physical access units. The main distinctive feature used for authentication is known as the *trabecular meshwork*, although other features are also used, such as furrows, freckles and the corona.

BEHAVIOURAL BIOMETRICS

- Voiceprint Recognition

Voiceprint recognition as the name would imply authenticates person by their vocal characteristics. The authentication can in principle be achieved both text dependently – where the user speaks a predefined word or sentence – and text independently where authentication is not dependent on the word(s) you speak, although, the latter is obviously a more difficult task to achieve successfully. Voiceprint recognition is similar to facial recognition and keystroke dynamics it that it can leverage existing hardware on the device, although some manufacturers do specify or provide a particular microphone that is calibrated with its authentication algorithm.

- Signature Recognition

This is achieved through using the distinctive aspects of a human signature to authenticate users. There are two underlying processes to signature recognition – static – where the completed signature is compared to a template version and authentication is given dependent on the comparison, or more comprehensively – dynamically – where behavioural components such as the speed, pressure and stroke order are also taken into account, hence making it less susceptible to forgery. The majority of signature-scan systems therefore use an electronic tablet that can record the dynamics of writing.

- Keystroke Dynamics

Keystroke dynamics is a technique that authenticates a person by the way in which they type on keyboards/keypads. The typical distinguishing characteristic is the latency between successive keystrokes. Similar to signature recognition, keystroke dynamics can be achieved using static and dynamic approaches, with the former being the easier. Static authentication involves the user entering a predefined keyword such as their username/password, whereas dynamic authentication is text

independent and will authenticate a user given any sequence of text. Since no additional hardware is required this has been a favoured technique, with much research on the subject since the 1980's (Gaines, 1980) but the performance of such a technique is comparatively weak against fingerprint and facial recognition systems, with currently only one commercially available product based on the static mode of authentication (Biopassword, 2002).

- Service Utilisation

This technique is achieved by monitoring the distinctive way in which a person interacts with a device. Measured factors could include the time and type of calls dialled (long distance, local, premium rate numbers for instance), SMS text messages sent to whom and when, and web pages visited over a period of time. The longer the period the more precise the technique becomes. The unique pattern(s) in a person's behaviour can be identified using a branch of artificial intelligence referred to as data mining (Singh et al., 2001). This is a comparatively new method of behavioural biometric and consequently has no commercial product to date.

The survey by Clarke et al. (2002a) also indicated that users wanted more security for their current second generation phones which in itself indicates user's awareness of security issues, and were prepared to use biometrics to achieve the desired level of security. Figure 2 illustrates user's responses towards some of the techniques previously described, considering their application to a mobile phone environment.

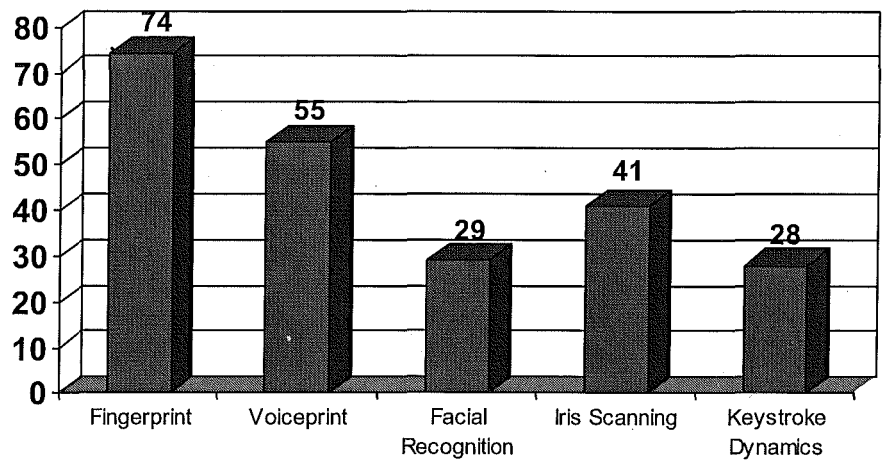


Figure 2 User's Biometric Preferences

The extent to which the biometrics previously described can be used within a mobile terminal device depends largely on the available hardware. It is unlikely, due mainly to cost, that many users will be willing to buy the additional hardware unless there are other real tangible benefits to be gained, such as a camera –which can be used for facial recognition but also take holiday pictures for instance. The only time where it would be conceivable for additional security-specific hardware purchases would be when the cost associated with the hardware is relatively small in comparison to the device to which it is protecting. This is likely to discount mobile handsets and PDAs as they are not likely to be expensive enough, but perhaps not laptops, where the upper boundary resides around $3700 ($6600 AUD). Otherwise it can be generally held true that the only biometric approaches available are those that can be easily (and cheaply) implemented on current devices. Typical biometric approaches that can be implemented on current mobile devices are given in table 1. This is by no means a definitive list as many devices differ in their hardware specifications. For instance some Acer laptops now have fingerprint recognition built into the system (Thornton, 2001) and some PDAs do not currently have the expandability to include a camera.

Mobile Phone	PDA	Laptop
➢ Voice Recognition via in-built microphone ➢ Keystroke Dynamics via scaled-down keyboard ➢ Facial Recognition via add-on or built-in camera ➢ Iris Recognition via add-on or built-in camera	➢ Voice Recognition via in-built microphone ➢ Facial Recognition via add-on camera ➢ Iris Recognition via add-on camera ➢ Signature Recognition via touch sensitive display	➢ Keystroke Dynamics via keyboard ➢ Fingerprint Scanner (via optional PCMCIA slot) ➢ Facial Recognition via in-built camera ➢ Iris Recognition via in-built camera

Table 1 Applicable Biometric for Mobile Devices

MOBILE BIOMETRICS IN PRACTICE

Keystroke dynamics is of particular interest as the approach has a number of advantages over other biometrics that make it useful as an authentication technique for mobile devices, mainly, the lack of additional hardware required and the ability to implement a solution completely transparently to the user, therefore resolving any issues of user inconvenience (the issues of convenience and intrusiveness are discussed in the following section). Although it is recognised that many PDAs do not have keyboards or keypads, a general market trend of late has seen the introduction of either add-on keyboards or scaled down versions (HP, 2002; HandSpring, 2002) to which keystroke dynamics can be applied. Of course no single biometric approach will encompass all mobile devices due to the differing hardware configurations, but the authentication mechanism proposed in this paper will take this into account.

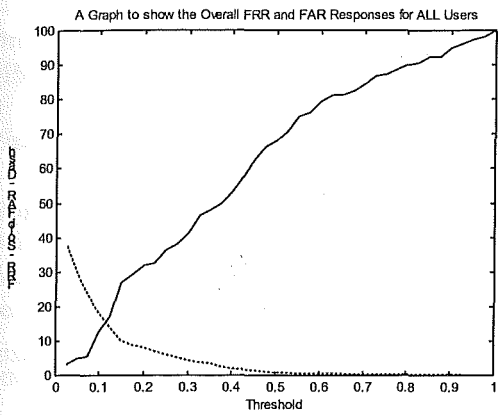
The history of keystroke dynamics dates back over twenty years with many research papers having been published, Joyce et al. (1990), Leggett et al. (1988) and Monrose et al. (1999) to name but a few. However, all studies to date with the exception of Ord (2000) have focussed on the ability to classify users on the basis of their interaction with a keyboard and not a keypad, as is common to mobile phones. To the authors best knowledge there have been no studies involving a mobile phone keypad – Ord's study used the numeric keypad from a computer keyboard, where the location and tactile differences are considered large enough to warrant an independent study. Thus a study was devised to investigate the feasibility of a keystroke dynamics technique on a mobile phone.

From the foundation Ord's study, a series of investigations were designed to examine the feasibility of using keystroke dynamics on a mobile handset (Clarke et al., 2002b). Three experiments were conducted, each involving a total of 16 participants:

1. the entry of a four digit number, analogous to the PINs used on current devices;
2. the entry of a series of varying telephone numbers;
3. the entry of a fixed telephone number.

The first and third investigations required the participants to enter the numeric keystroke sample thirty times, with twenty samples then being used to create a reference profile, and the remaining ten for subsequent testing. The second investigation required a larger number of samples due to the changing nature of the input string, and thus the need to train the authentication system more accurately. Fifty samples were taken, with thirty for training and twenty for testing.

Previous studies have shown neural networks to provide an effective foundation for keystroke analysis (Ord, 2000; Cho et al., 2000) and they have consequently been used in these investigations. The neural network structure is constructed on the feed-forward back-propagation network (Bishop, 1995), best exemplified for pattern recognition techniques.



Investigation	FAR (%)	FRR (%)	EER (%)
PIN Code	18.1	12.5	15
Varying Telephone	36.3	24.3	32
Fixed Telephone	16	15	15

Table 2 Keystroke Dynamics Results

Figure 3 Keystroke Dynamics Performance Chart

The results demonstrate the potential to distinguish authorised users from impostors, although arguably not to any great accuracy. However, the experimental procedure used in this study was performed under controlled conditions, with users all entering the same input data - a condition that is unlikely in the real world. Additionally, the design, and implementation of the neural network used for classification was primitive and un-optimised. Continuation of the study beyond this feasibility stage requires variables such as pre-processing, generalisation, network sensitivity and network configuration to be considered and analysed.

Further development of the technique will also consider other forms of user interaction with mobile handsets, in order to attempt to profile behaviour in different contexts. For instance, the way in which someone types when entering an SMS message is likely to be different to the way in which they enter a telephone number. Some users will use certain applications or functionality on the phone more often than others; will dial certain number more than others; and equally as important will not use or dial certain people or services. All of these factors could potentially be used as discriminating characteristics, leading to a stronger overall verification technique.

CONCLUSIONS

Mobile devices are going through an evolutionary period with the combined ability to have high computer processing on small handheld devices, and the formidable success of the mobile phone industry. Users are no longer chained to their desks and mobility has become an important factor in many people's life. This has left an increasing security problem generally, with a major issue being authentication.

The current form of authentication is a very cheap solution but suffers from a number of inherent weaknesses, such as the lack of and improper use of passwords and PINs. Biometrics are amongst the most powerful authentication tools as they are based on a unique human characteristic.

Biometrics' on mobile devices are also an effective tool for non-intrusive authentication, as different approaches can be implemented whilst the user is interacting with the device. In the context of a mobile phone, voice recognition can be used to authenticate a user whilst they are speaking on the phone, keystroke dynamics whilst they are typing SMS messages and facial recognition when they use video conferencing facilities. Thus a hybrid non-intrusive authentication mechanism utilising the

available biometrics on each mobile device as the underlying authenticator would provide a transparent and secure solution.

REFERENCES

Ashbourn, J. (2000). Biometric. Advanced Identity Verification. The Complete Guide. Springer.

BBC. (2002). MOD Loses 600 Laptops. BBC News Online http://news.bbc.co.uk/hi/english/uk/newsid_1757000/1757792.stm

Biopassword (2002). www.biopassword.com

Bishop, M. (1995). Neural Networks for Pattern Recognition. Oxford University Press.

Broomfield, S. (2000). It's Not Just a Laptop Anymore!. Information Impacts Magazine. www.cisp.org/imp/february_2000/broomfield/02_00broomfield.htm.

Cho, S., Han, C., Han, D., Kim, H. (2000). Web Based Keystroke Dynamics Identity Verification using Neural Networks. Journal of Organisational Computing & Electronic Commerce, Vol.10, No.4, pp.295-307.

Clarke, N., Furnell, S., Rodwell, P., Reynolds, P. (2002a). Acceptance of Subscriber Authentication for Mobile Telephony Devices. Computers & Security, Vol.21, No.3, pp.220-228.

Clarke, N., Furnell, S., Lines, B., Reynolds, P. (2002b). Subscriber Authentication for Mobile Phones using Keystroke Dynamics. Proceedings of the Third International Network Conference (INC2002), Plymouth, UK, 16-18 July 2002. pp347-356.

Cope, B. (1990). Biometric Systems of Access Control. Electrotechnology, April/May: 71-74.

Gaines, R., Lisowksi, W., Press, S., Shapiro, N. (1980). Authentication by keystroke timing: Some Preliminary Results. Rand Report R-256-NSF. Rand Corporation, Santa Monica, CA.

Gibson, B. (2001). Apple Slips to 8th in US laptop sales. MacCentral Online. http://maccentral.macworld.com

Giussani, B. (2001). Roam – Making Sense of the Wireless Internet. Random House Business Books, London.

HandSpring. (2002). Treo 270. HandSpring. www.handspring.com

Harrington, V., Mayhew P. (2001). Home Office Research Study 235: Mobile Phone Theft. Crown Copyright.

Harrison, L. (2001). Iris Recognition is Best Biometric. The Register. www.theregus.com
Joyce R., Gupta, G. (1990). Identity Authentication Based on Keystroke Latencies. Communications of the ACM. Vol. 39; pp 168-176.

Leggett, J., Williams, G. (1988). Verifying Identity via Keystroke Characteristics. International Journal of Man-Machine Studies, 28.

Leyden, J. (2002). PDAs Make Easy Pickings for Data Thieves. The Register. www.the register.co.uk/content/54/25478.html

Monrose, F., Reiter, M., Wetzel, S. (1999). Password Hardening Based on Keystroke Dynamics. Proceedings of the 6th ACM Computer and Communication Security Conference.

Nanavati, S., Thieme, M., Nanavati, R. (2002). Biometrics. Identity Verification in a Networked World. John Wiley & Sons.

Ord, T. (1999). User Authentication Using Keystroke Analysis with a Numerical Keypad Approach. MSc Thesis, University Of Plymouth, UK.

Richardson, T. (2002). PDA Shipment Growth Slows. The Register. www.theregister.co.uk.

Singh, H., Furnell, S.M., Lines, B. and Dowland, P.S. (2001). Investigating and Evaluating Behavioural Profiling and Intrusion Detection Using Data Mining. Proceedings of International Workshop on Mathematical Methods, Models and Architectures for Computer Networks Security, St. Petersburg, Russia, 21-23 May, 2001.

Smith, R. (2002). Authentication. From Passwords to Public Keys. Addison Wesley

Thornton, C. (2001). Fast laptop includes built-in fingerprint reader. PC World.com. www.pcworld.com/news.

HP. (2002). Targus iPAQ™ Bundle. HP. http://athome.compaq.com/store/default.asp?page=optionCategories&SuperCategoryID=97

UMTS Forum (2002). Long Term Potential Remains High For 3G Mobile Data Services. www.umts-forum.org/reports.html. pp 5.

An Investigation Into the Application of Defence In Depth Theory to Electronic Information Protection

Andrew J Lester[1] and Clifton L Smith[2]

Security and Applied Technology Research (SATR) Group
School of Engineering and Mathematics
Edith Cowan University
Joondalup Campus, Western Australia

[1]Intelligent Services
PO Box 3126
Joondalup, Western Australia

[2]Machine Vision Research Group
Department of Electrical and Electronic Engineering
Nottingham Trent University
Nottingham, England

ABSTRACT

This paper discusses an investigation into the application of traditional Defence in Depth theory to digital electronic information protection. Defence in Depth is firstly discussed in a physical security context, where deterrence, detection, delay and response are shown to be achieved by psychological, electronic, physical and procedural barriers. The Electronic Information Attack Model is then proposed, which comprises a hierarchical structure defining different aspects of electronic information and ways of attacking its confidentiality, integrity and availability. The final component then proposes that the four Defence in Depth functions can provide electronic information protection by layering barriers at various levels in the Electronic Information Attack Model.

Keywords: *digital electronic information, Defence in Depth, Electronic Information Attack Model, information security, barriers*

INTRODUCTION

This paper discusses an investigation into the application of the traditional physical security concept of Defence in Depth to the modern asset of digital electronic information. While Defence in Depth principles have been employed for centuries in protecting facilities, there are now opportunities to extend their use to cover electronic information as it becomes increasingly valued and ubiquitous as an organisational resource. Although Defence in Depth is often inherently employed for electronic information protection in practice, this paper formalises the theory underlying this approach using a new model describing the structured relationship between information, threats, and protective measures.

The theoretical models in this paper have been limited to the protection of digital electronic information against deliberate attacks carried out by human threats. Using the same principles, however, there are opportunities to expand the investigation to incorporate a wider scope with maturation of the theory.

The first section of this paper examines traditional Defence in Depth theory in terms of its application to the protection of facilities and assets in physical form. The second section proposes the Electronic Information Attack Model, which has been developed to assist in analysing different ways in which electronic information can be attacked. The third section then examines the application Defence in

Depth in the context of electronic information protection by combining the two models that have been addressed.

DEFENCE IN DEPTH THEORY

In a physical security context, the principle of Defence in Depth specifies a series of barriers to surround the assets being protected (Lester et al 1999). These barriers resist an intruder's access to a facility's assets, and have the functions of:

- *Deterrence*: Discourages and warns against unauthorised access;
- *Detection*: If an intrusion has commenced, the presence of perpetrators needs to be registered so that an appropriate response can be initiated;
- *Delay*: Sufficient physical barriers should slow the intruder's progress so that enough time is provided for a positive response; and
- *Response*: An appropriate response to intrusion will be the apprehension of the perpetrators and the neutralisation of the attack.

The barriers used in a facility's Defence in Depth strategy are not necessarily physical or material in form, but instead may be considered according to the functions they provide (Smith et al 1999):

- *Psychological barriers*: These have the function of deterring opportunist intruders, and may include signage, low fences, glare security lighting, and closed circuit television (CCTV) systems.
- *Electronic barriers*: These have the function of detecting intruders when entry to the facility has been made. They encompass the entire range of detection technology and include access control systems, biometric identification (Smith 1997), intrusion detection systems, and CCTV systems.
- *Physical barriers*: These have the function of delaying the progress of a determined intruder, and can include walls, fences, locks, safes and strengthened glass.
- *Procedural barriers*: These provide the operational or management function of the Defence in Depth strategy, and can include the wearing of identification badges, the escorting of visitors in the facility, and appropriate procedures for responding to an intrusion.

The concept of multiple barriers of various types is central to the Defence in Depth strategy for effective protection of assets from intruders.

ELECTRONIC INFORMATION ATTACK MODEL

The approach taken in this investigation was to firstly develop what has been termed the *Electronic Information Attack Model* (EIAM). This model categorises information according to its attributes, and uses the resulting structure to examine possible ways by which it can be compromised. Such an approach provides a platform that can accommodate protective barriers, thereby facilitating the application of Defence in Depth.

The categorisation has involved organising information into a hierarchy of levels, with the upper section considering information itself, and the lower areas considering different methods of attack. Figure 1 shows the structure of the EIAM, which contains the following levels:

Level 0: Information
Level 1: Category of information
Level 2: State of digital electronic information
Level 3: Means of accessing digital electronic information in each state
Level 4: Means of attacking digital electronic information for each mode of access
Level 5: Specific means of attacking confidentiality, integrity and availability

The transition down this hierarchy shows a shift in emphasis from the information assets themselves to the ways in which threats can act against these assets in the form of an attack. The discussion to follow outlines each of these levels.

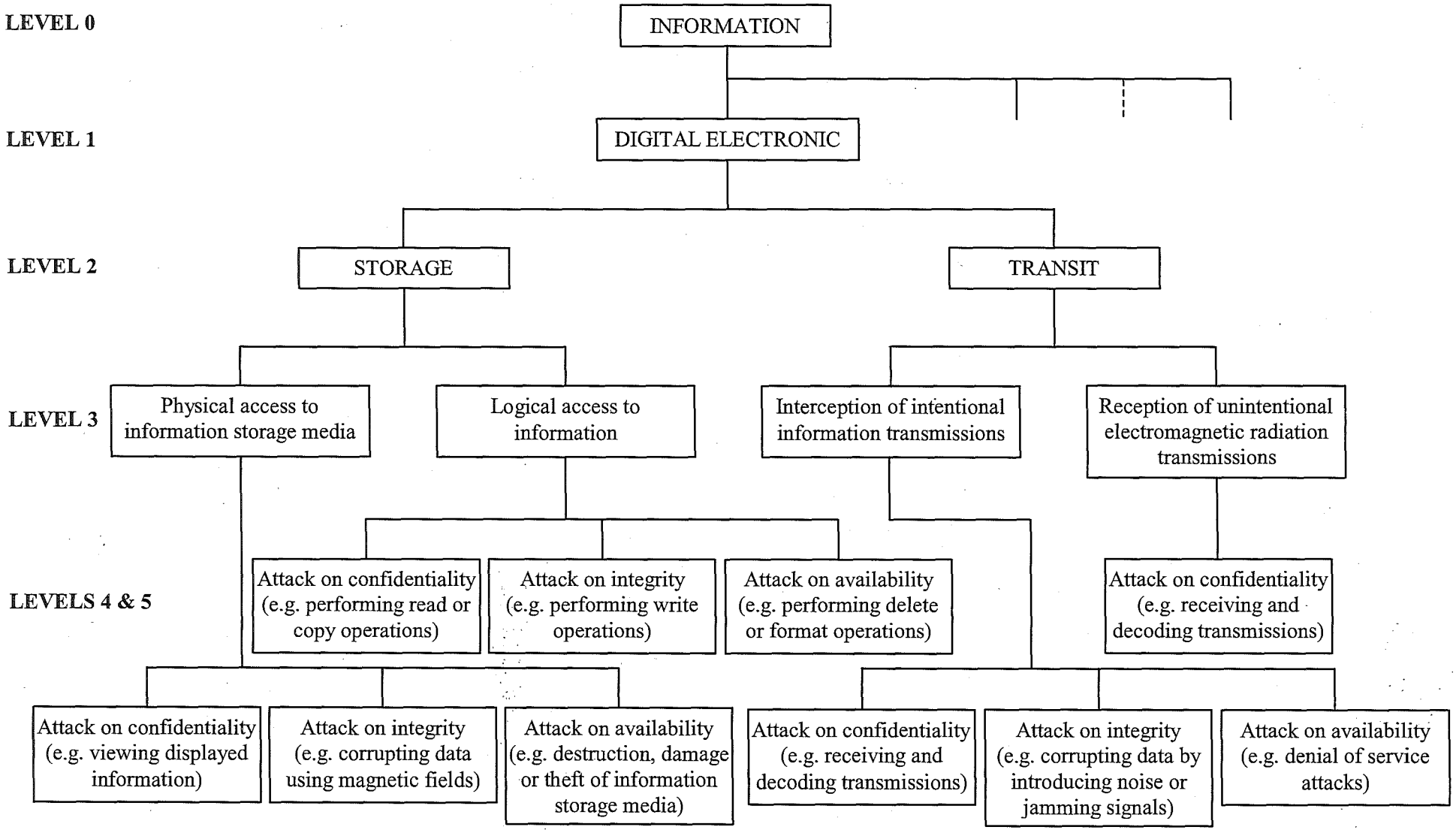


Figure 1: Structure of the Electronic Information Attack Model

Level 0: Information

Organisations possess a combination of personnel, information, property and intangible assets (Lester 2001), and this level defines the EIAM context by focussing on information as one kind of asset to be protected. Level 0 incorporates all information as a whole, regardless of the form in which it exists.

Level 1: Category of Information

Information can be broken into various categories based on the ways in which it is stored, transferred and used. One particular form is digital electronic information, some examples of which include computer system data, smart card information, and digital telecommunications transmissions. This level delimits the scope of the EIAM, although by defining other specific categories at this level, the EIAM can be expanded to become a more generic Information Attack Model. This could then be used to investigate the application of Defence in Depth in the context of information security as a whole.

Level 2: State of Digital Electronic Information

Digital electronic information always exists in one of two states (Vuori 1997):

1. Storage; or
2. Transit.

Storage implies that information is stationary, and therefore incorporates instances in which it is held on any media or viewable on any visual display unit. Some examples include information on hard or floppy disc drives, compact discs, in random access memory, and that being displayed on computer monitors.

Transit implies that information is travelling from one point to another, which usually involves the transmission of signals. Some examples include the transfer of data via computer networks, mobile wireless communication, and also unintentional electromagnetic radiation emanations from electronic information handling equipment and infrastructure.

Level 3: Means of Accessing Digital Electronic Information in Each State

Depending on whether information is in storage or transit, there are different ways in which it can be accessed by an attacker. The term 'access' is used in this case to describe any situation in which an attacker, or one of their instruments, makes contact with either information itself, or devices or signals with which it is related.

When in storage, there are two means of accessing information:

1. Making physical contact with the information storage media (which includes equipment used for displaying or processing information in this model); or
2. Making logical contact with the information itself.

When in transit, there are two means of accessing information:

1. Receiving, intercepting, or otherwise interfering with signals transmitted intentionally between transmitting and receiving devices; or
2. Receiving and reconstructing information from unintentional electromagnetic radiation signals emanating from electronic devices or cabling (Defence Signals Directorate 2000).

These four methods each provide different opportunities for information to be compromised, since the specific attack will depend on how the attacker has gained access.

Level 4: Means of Attacking Digital Electronic Information for Each Mode of Access

Having defined ways in which information, equipment and signals can be accessed, this level considers the different elements that can be attacked once access is gained. Depending on the context and type of information, there are three requirements for it to remain secure (Pangalos et al 1994):

1. Its confidentiality is maintained: meaning that the information is only known and accessible by authorised personnel;
2. Its integrity is maintained: meaning that the information is complete and has not undergone any unauthorised or undesirable modification; and
3. Its availability is maintained: meaning that the information is accessible whenever required.

Once an attacker has accessed information, there is the potential for one or more of these requirements to be compromised, thereby compromising the information's security. It should be noted that in some cases, confidentiality may not be necessary, however information must always have integrity and availability.

Level 5: Specific Means of Attacking Confidentiality, Integrity and Availability

Depending on which information security requirements are being attacked, there are specific means by which they may be compromised, some examples of which are shown in Figure 1.

INTEGRATING DEFENCE IN DEPTH WITH THE ELECTRONIC INFORMATION ATTACK MODEL

A fundamental Defence in Depth principle is the layering of barriers between an asset and corresponding threat(s). In most cases, a single barrier cannot provide all four Defence in Depth functions of deterrence, detection, delay and response, and so the combination of multiple contributing barriers is necessary (National Crime Prevention Institute 1986). This can be achieved in the context of the EIAM by applying one or protective measures at one or more of its levels.

Depending on a barrier's purpose, it will be suited to a particular level or region of the EIAM. Those barriers that defend against a specific type of attack, such as anti-virus software, will be applied at the lower levels, while more generic barriers that offer protection in a wider range of situations against multiple types of attack, such as an information security policy, would be situated at the upper levels.

The application of barriers at various levels in the EIAM is analogous to providing rings of defence around a facility's assets in a physical security context. When analysing the level of protection provided by this Defence in Depth approach, however, the number of layers must be interpreted in terms of the specific attacks being considered. This is because not all barriers in a total information security scheme will be effective against all means of compromising information.

The translation of the four traditional Defence in Depth functions to electronic information protection is discussed as follows:

a. *Deterrence*: This function is performed by any barriers that discourage a potential attacker from carrying out their attack (for example, Crime Prevention Through Environmental Design techniques (Crowe 1991) applied at level 3 to protect against physical access to information storage media).
b. *Detection*: This function is performed by any barriers that either:
    1. Detect the progress of an attack (for example, network intrusion detection applied at level 3 to protect against logical access to information, or human motion sensing applied at level 3 to protect against physical access to information storage media); or

2. Detect information compromise (for example, error detection applied at level 4 to identify data corruption during transmission, or procedures applied at level 2 to determine whether information has been compromised while in storage).

c. *Delay*: This function is performed by any barriers that increase the amount of time required to perform an attack. The delay can be of:
   1. Finite duration (for example, encryption applied at level 1); or
   2. Infinite duration, in which a particular attack is completely denied (for example, physically isolating a PC from a computer network at level 3 to avoid external parties from logically accessing information, or using fibre optic cables for data transmission at level 2 to avoid interception, electromagnetic emanations or noise corruption).

d. *Response*: This function is performed by any barriers that neutralise an attack or the compromise of information, which can involve either:
   1. Disabling an attack whilst in progress (for example, a human response force applied at level 3 to apprehend an intruder attempting physical access to information storage media); or
   2. Restoring information security following its compromise (for example, applying error correction at level 4 to return corrupted data to its original state after transmission, or restoring information from backups at level 5 following destruction of information storage media).

As is evident, the application of Defence in Depth to the protection of electronic information introduces many unique characteristics that make it more complex than its application in physical security.

The four types of barriers in the traditional Defence in Depth model are also applicable to electronic information protection. A brief analysis of the specific ways in which they achieve each Defence in Depth function is as follows:

e. *Psychological barriers*: Provide deterrence alone (for example, warning signs applied at level 3 to protect against unauthorised physical access to information storage media, or warning messages applied at level 3 to protect against unauthorised logical access to information).

f. *Physical barriers*: Provide a delay of either finite or infinite time (for example, fences and security containers applied at level 3 to protect against access to information storage media, or TEMPEST certified equipment applied at level 3 to avoid electromagnetic radiation emanations).

g. *Electronic barriers*: When considering Defence in Depth for electronic information protection, it is appropriate to divide electronic barriers into two types:
   1. Physical, which are those barriers that have an electronic basis yet exist in physical form (for example, closed circuit television systems and human motion sensors applied at level 3 to protect against physical access to information storage media); or
   2. Logical, which are those barriers which are intangible and are integrated with information coding, processing and handling (for example, encryption and digital watermarks applied at level 1, anti-virus software applied at level 4 for protection during storage, or spread spectrum wireless transmission applied at level 4 for protection during transit).

   In most cases, electronic (physical) barriers can provide only deterrence and detection, however it is possible for various electronic (logical) barriers to provide deterrence, detection, delay or response functions depending on their nature and implementation.

h. *Procedural barriers*: Provide deterrence, detection (for example, security education, training and awareness at level 1), delay (for example, separation of duties at level 1, or requirements for multiple persons to enter passwords at level 3 for logical access to information) and response (such as a guard force applied at level 3, or offsite backups applied at level 2 to return stored information). They can also contribute towards the implementation of other barriers (such as an information security policy).

Therefore, in an electronic information protection context, each barrier type is not primarily concerned with only one Defence in Depth function (as discussed in physical security applications), but instead, there is a greater diversification in the roles of each barrier type.

CONCLUSION

This paper has presented an investigation into the application of Defence in Depth theory to digital electronic information protection. The first section has covered Defence in Depth in a physical security context, with an emphasis on barriers and the security functions they provide. The second section has then presented a new model for examining the various ways human threats may attack electronic information. The third section has looked at the ways in which Defence in Depth can be used for electronic information security by employing barriers to protect against these attacks. It has also outlined some of the characteristics of Defence in Depth that are unique to its application in electronic information protection.

This investigation shows that Defence in Depth can be used for electronic information protection, however its application is more complex than in a physical security scheme. Although this paper represents an advancement in the development of electronic information security theory, further investigation and development is required to obtain a comprehensive and exhaustive model. In addition, by removing this paper's scope limitations, it will be possible to obtain a more generic and widely applicable model to further assist in the practical implementation of security measures for electronic information protection.

REFERENCES

Crowe, T.D. (1991). *Crime Prevention Through Environmental Design*, Butterworth-Heinemann, Boston.

Defence Signals Directorate. (2000). *Australian Communications – Electronic Security Instruction 33*, Commonwealth of Australia, Canberra.

Lester, A.J. (2001). *Asset-Barrier Integration*. Unpublished manuscript, Edith Cowan University, Perth.

Lester, A.J. and Smith, C.L. (1999). Analyses of Performance of Volumetric Intrusion Detection Technologies. *Proceedings of the 33rd Annual International Carnahan Conference on Security Technology*, pp. 101-111.

National Crime Prevention Institute. (1986). *Understanding Crime Prevention*, Butterworth-Heinemann, Boston.

Pangalos, G., Frangakis, C. and Katsikas, S. (1994, Fall). Implementing Database Systems Security. *Computer Security Journal* **X**(2), pp. 73-86.

Smith, C.L. (1997). Identification by biometric systems. *Focus International*, **3**(2), pp. 33-40.

Smith, C.L. and Robinson, M. (1999). The Understanding of Security Technology and its Applications. *Proceedings of the 33rd Annual International Carnahan Conference on Security Technology*, pp. 26-37.

Vuori, T. (1997). Lecture Notes: Computer Security. Unpublished manuscript, Edith Cowan University, Perth.

Managing Electronic Banking Risks: An Overview

Steven Li[1] and Zhongwei Zhang[2]

[1]*School of Economics and Finance*
Queensland University of Technology
Brisbane, Australia
E-mail: s.li@qut.edu.au

[2]*Department of Mathematics and Computing*
University of Southern Queensland
Toowoomba, Australia
E-mail: zhongwei@usq.edu.au

ABSTRACT

This paper is concerned with e-banking risks. The general e-banking risks are identified and some risk management solutions are proposed and analysed. In particular, we show how to assess e-banking risk exposures and discuss the need of hedging and insuring e-banking risks.

Key words: e-banking, risks, risk map, risk management, insurance and hedging

INTRODUCTION

Banks have been delivering electronic services to consumer and business remotely for years. Electronic funds transfer, including small payments and corporate cash management systems, as well as publicly accessible automated machines for currency withdrawal and retail account management, are already global fixtures. However the widely acceptance of electronic delivery channels, especially the Internet, for banking products and services provides new business opportunities for banks and service benefits for their customers.

Electronic banking (e-banking) refers to the provision of retail and small value banking products and services through electronic channels as well as large value electronic payments and other wholesale banking services delivered electronically. These include traditional activities such as accessing financial information, obtaining loans and opening deposit accounts, as well as relative new products and services such as electronic payment services, personalized financial "portals", account aggregation and business-to-business market places and exchanges.

E-banking is a segment of electronic commerce ("e-commerce"), which, in turn, encompasses all types of business transacted through electronic networks. In other words, e-banking is not a banking product; rather it describes the way transactions are conducted. A summary of all types of e-banking is presented in the graph below.

"PC banking" is the term used for banking business transacted from a customer's PC. The exchange of data involved, for instance, in the transmission of orders for credit transfers to the bank is effected via phone lines (either analog—by modem—or using an ISDN adapter). Basically, there are two types of PC banking: online banking and Internet banking. Unlike online banking in which bank transactions are conducted within closed networks, Internet banking permits the customer to conduct transactions from any terminal with access to the Internet.

Mobile banking is a vivid example of how the lines between the various forms of e-banking are becoming blurred. Thanks to new transmission technologies such as WAP (wireless application protocol), portable digital assistants (PDAs) or small hand-held PCs are providing bank customers with access to the Internet and thus paving the way to mobile banking. As a result, Internet banking has become more than just a form of PC-banking.

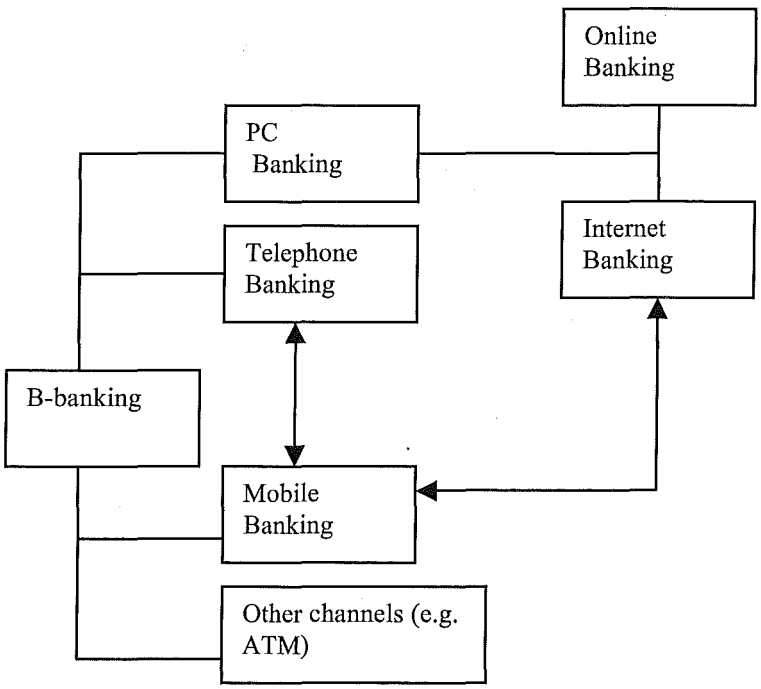


Figure 1 Source: Deutsche Bundesbank

In recent years, Internet-banking has been a focus for e-banking and it has made a great progress so far. However, there have been numerous incidents that raised concerns about e-banking risks. To name a few (cf. Hutchinson and Warren, 2001):

Barclays, which claims to be the UK's largest online bank, had to take down its website at the end of July, when customers were served the bank statements of other clients (Knight 2000a & 2000b).

Western Union, part of Atlanta-based electronic payments giant First Data Corp. implemented a new Internet based person-person payment service but the site was hacked during a routine maintenance operation that erroneously left parts of the sites exposed (Creed 2000).A report by Association Press detailed how hackers broke in and copied the credit card and debit card details of 15700 Western Union customers who use the site to transfer funds across the Internet (Gutterman 2000).

Notwithstanding the significant benefits of technological innovation, the rapid development of e-banking capabilities carries benefits as well as risks and it is important that these risks are recognized and managed by banking institutions in a prudent manner. In this paper, we shall review risk management challenges facing e-banking as well as the types of e-banking risks. After gaining a thorough understanding on e-banking risks, then we consider how to manage e-banking risks, where some possible new methodologies are discussed. Finally, some general principles on e-banking risk management are given.

THE NATURE OF E-BANKING AND ITS CHALLENGES FOR RISK MANAGEMENT

The world is becoming increasingly open as a result of the Internet and the World Wide Web (WWW). Due to the rapid development in telecommunication and computing technology, electronic

banking has been gaining ground around the globe. For example, a recent survey by online financial services provider Egg indicating that British Internet users are becoming more likely to use financial services online (Anonymous, 2002). This offers banking institutions a new frontier of opportunities and challenges further augmenting competition in the global banking market.

The rapid development of e-banking is also largely due to the benefits it brought to customers and banks. It offers great convenience and flexibility to customers. For example, e-banking can offer access to transactions, account balance and information anywhere and anytime, which is in stark contrast with the traditional branch banking during business ours.

From a bank's point of view, a sound business rationale does exist for e-banking (Robison 2000). E-banking requires high initial set-up costs (both technological and marketing) with savings following later. Once marketing and set-up costs have been incurred, transaction costs (admittedly, excluding the cost of customer support) appear much lower for e-banking, especially in high-wage economies (Table 1 below).

	United States	India
Physical branch	100	100
Postal	..	40
Telephone	50	18
ATM	27	18
PC dial –up	8	na
Internet	1	12
Sources: Sato et al (2001)		

Table 1: Relative costs of banking transactions

The more transactions that can be converted to electronic form, the more money will be saved. The cost of an electronic transaction is dramatically less when done online, and customers do most of the work themselves. Online banking has also the potential to solidify and extend a bank's relationship with its customers because it brings banking services directly to a customers home or office. The more services a customer accepts, the more likely that customer will stay loyal to the bank. Besides, online services are a must for banks that have to compete with a growing number of services from other financial institutions, investment concerns and insurance companies.

The opinion that traditional banks were "dinosars" that the Internet would drive to extinction is no longer widely held. A study comparing new Internet-only banks with a peer group of new branch banks by De Young (2001) shows the Internet-only banks have been substantially less profitable. They generate lower business volumes and any savings generated by lower physical overheads appear to be offset by other types of non-interest expenditures, notably marketing to attract new customers. However, Internet-only banking could eventually prove to be a viable business model: De Young finds that profitability improves more quickly over time for the Internet only start-ups and they may benefit more from gaining experience and be better placed to realise economies of scale than their peers.

Most researchers and practitioners believe that dis-intermediation is unlikely to occur and financial intermediation is still essential in the age of the Internet (Lin Geng & Whinston, 2001).

Security issues and prevailing attitudes that financial products are too complex to select without face-to-face guidance from a professional have kept the average consumer away from Internet banking, while online stock trading and insurance have had much faster growth (Alstad 2002).

The fundamental characteristics of e-banking (and e-commerce more generally) posed a number of risk management challenges (cf. BIS, 2001):

- The speed of change relating to technological and customer service innovation in e-banking is unprecedented. Historically, new banking applications were implemented over relatively long periods of time and only after in-depth testing. Today, however, banks are experiencing competitive pressure to roll out new business applications in very compressed time frames—often only a few months from concept to production. This competition intensifies the management challenge to ensure that adequate strategic assessment, risk analysis and security reviews are conducted prior to implementing new e-banking applications.
- Transactional e-banking web sites and associated retail and whole sale business applications are typically integrated as much as possible with legacy computer systems to allow more straight-through automated processing reduces opportunities for human error and fraud inherent in manual processes, but it also increases dependence on sound system design and architecture as well as system interoperability and operational scalability.
- E-banking increases banks' dependence on information technology, thereby increasing the technical complexity of many operational and security issues and furthering a trend towards more partnerships, alliances and outsourcing arrangements with third parties, many of whom are unregulated. This development has been leading to the reaction of new business models involving banks and nonbank entities, such as internet service providers, telecommunication companies, and other technology firms.
- The Internet is ubiquitous and global by nature. It is an open network accessible from anywhere in the world by unknown parties, with routing of messages through unknown locations and via fast evolving wireless devices. Therefore, it significantly magnifies the importance of security controls, customer authentication techniques, data protections and audit trail procedures, and customer privacy standards.

In order to manage risks, a firm must know what risks it faces and how big they are. Consequently, the firm must implement a 'system' for measuring risk. This general rule does apply to managing e-banking risks. In this paper, we attempt to identify the key risks in e-banking and propose some solutions for managing these risks. In particular, we shed some lights on the further development of hedging and insuring e-banking risks.

E-BANKING RISKS

Although IT innovations has afforded new business opportunities to banks, the risks involved in e-banking also pose enormous challenges for banks as illustrated in the previous section. Because of the rapid changes in information technology, no list of risks can be exhaustive. Specific risks faced by banks engaged in electronic banking can be grouped according to risk categories discussed in Basle Committee risk management documents, in this sense, the risks are not new

While the basic types of risks generated by e-banking are not new, the specific ways in which some of the risks arise, as well as the magnitude of their impact on banks, may be new for banks and supervisors. E-banking is causing a shift in the weighting of existing risk categories towards those risks arising from the increased use of IT. According to a recent study by Deutsche Bundesbank (2000), the major types of risks involved with e-banking can be categorised follows.

Strategic risk

Strategic risks result from bad business decisions taken by management. Specifically, the danger of not being able to keep up with rival technologies is the source of greatest strategic risk. Technology is so important for e-banking operations that there is a correspondingly great need to invest in new technologies.

The rapid pace of innovation in e-commerce is requiring banks to make e-banking strategy decisions as quickly and intelligently as possible, since technological innovations or changes in customer tastes caused by "waves of fashion" often make radical adjustment inevitable. Frequently there is no way of

predicting which technology and which terminals (e.g. mobile phones, television sets, PDAs) will ultimately prevail.

Operational risk

Operational risk in the narrower sense encompasses all risks originating directly in business operations. Important sources of operations risk include technical malfunctions or human error, It problems, fraud and inadequate organisation structures. If operational risk is not managed efficiently, tis could result not only in financial losses but also in disruptions in banking operations (e.g. a call centre can not be reached, or a host system is down). Operational risk is by no means new, however, the increasing use of IT in recent years has been making it more and more conspicuous.

Legal Risk

Legal risk derives from the fact that laws governing the validity and enforceability of electronically concluded agreements are only now being drafted in may countries.

Generally speaking, the regulatory approaches continue to vary from one country to the other. The resultant uncertainty is heightened in those cases where cross-border transactions involve countries where the credit situation has no physical presence. Additionally, a lack of familiarity with foreign systems leads to risks regarding consumer and data protection issues.

Reputational risk

Banking business is especially sensitive to fluctuations in confidence. Therefore, reputational risk, particularly in a relatively new field of business, represents a special challenge for banks. Customers' confidence in their bank can be shaken if the bank is not able to provide secure and trouble-free e-banking services. The same is true if services such as responding to inquiries or processing orders are not performed at the speed that customers have come to expect in the "Internet Age".

Systemic risks

On the downside, banking supervisors must analyse not only individual risks but also macro-prudential implications (systemic risk) of e-banking. There is no disputing in the fact that e-banking has changed the risk structure of the banking sector, by, for instance, increasing operational risk. Moreover, banking supervisors will have to keep up with torrid pace of innovation on the Net.

On a more positive note, e-banking opens up new sources of profits to the banking sector, since banks can achieve cross-selling effects by, for instance, joining forces with non-banks. In addition, the competition-enhancing effect of e-banking may encourage a sort of house cleaning as far as the structure goes by, for instance, forcing a bank to streamline its branch office network. On the whole, e-banking is becoming a more and more important segment for macro-prudential analysis, and should be given closer scrutiny in future; however, at present this segment would not appear to pose any exceptional systemic risk.

RISK MANGEMENT

For an increasing number of banks there maybe a strategic reason for engaging in electronic banking. In addition, greater use of electronic banking may increase the efficiency of the banking and payment system, benefiting consumers and merchants. At the same time, as the proceeding discussion indicates, there are risks for banks engaging in electronic banking. Risks must be balanced against benefits; banks must be able to manage and control risks and absorb any related losses if necessary. Risks for e-banking should also be evaluated in the context of other risks that the bank faces.

The rapid pace of technological process innovation is likely to change the nature and scope of risks banks face in e-banking. A systematic risk management process proposed by BIS (1998) consists of three basic elements of *assessing risks, controlling risk exposure and monitoring risks*. It is essential that banks have a comprehensive risk management process in place that is subject to appropriate oversight by the board of directors and senior management.

RISK ASSESSMENT

Assessing risks is an ongoing process. A bank may engage in rigorous analytic process to identify risks and, where possible to quantify them. Risk assessment involves a combination of a variety of formal and informal methods. It is practised in a variety of areas, by individuals with a wide spectrum of skills. Its goal is to determine the probabilities and impacts of various events. This data then can be used to provide guidance on the proper management of risks.

A significant risk assessment tool is a simple graph known as a risk map. A risk map is a two-dimensional chart that shows the probabilities on one axis and the consequences (losses) on the other. An application in the e-commerce setting, which includes E-banking, is given in (Li, 2001).

The general risk assessment rules can be applied to e-banking. In assessing the risks related to e-banking, we need first to estimate the consequences associated with the risks. This may not be too hard to estimate in practice. Then we need to estimate the probabilities related to each risk. In practice, this may be hard and may not be accurate enough due to the shortage of historical data and comparison data. For a bank with sufficient historical data, we can use the historical data to predict the probabilities of the loss due to, for example, hackers in a specific period. Of course, such probability may be evolving overtime due to technology progress or changes in legal environment.

Armed with the information presented on the risk map, risk events can be modelled using simulation. Simulations are computer programs that use probability and loss information as inputs. They can be used to develop and test various risk management options in a realistic yet controlled environment. It is in this sense that the probability/loss analysis precedes that development of useful holistic risk management techniques. We need to know what we are up against before we can develop workable solutions.

MANAGING AND CONTROL RISKS

Having made an assessment of risks and its tolerance, bank management should take steps to manage and control risks. This phase of a risk management process includes activities such as implementing security policies and measures, coordinating internal communication, evaluating and upgrading products and services, implementing measures to ensure that outsourcing risks are controlled and managed, providing disclosure and customer education and developing contingency plans.

Banks are used to and good at managing financial risks such as interest rate risk and foreign currency risk, which are often controlled by hedging. Below we attempt to shed some light on how banks can handle the e-banking risks effectively.

Security policies and measures

Security is the combination of systems, applications, and internal controls used to safeguard the integrity, authenticity, and confidentiality of data and operating processes. Proper security relies on the development and implementation of adequate security policies and security measures for processes within the bank, and for communication between the bank and external parties. Security policies and measures can limit the risk of external and internal attacks on electronic banking systems, as well as the reputational risk arising from security breaches.

A security policy states management's intentions to support information security and provides an explanation of the bank's security organization. It also establishes guidelines that define the bank's security risk tolerance. Security measures are combinations of hardware and software tools, and personnel management, that contribute to building secure systems and operations (BIS, 1998). Security policies can be used to screen out some avoidable risks related to e-banking. The following aspects should be considered in drawing up appropriate policies for managing e-banking risks.

- Internal communication
  To ensure adequate internal communication, all policies and procedures should be provided in writing. In addition, senior management should adopt a corporate policy of ongoing education and upgrading of skills and knowledge in order to limit operational risks arising from lack of staff and management expertise.
- Evaluating and upgrading
  Before introducing products and services, they should be tested on a widespread basis. Pilot programs and prototype s can be helpful in developing new applications.
- Outsourcing
  When outsourcing, banks management should evaluate the ability of the service provider to maintain the same level of security as though the activities were conducted in house, through the review of the service provider's policies and procedures aimed as protecting sensitive data.
- Disclosure and custom education
  Disclosure and explanations about the nature of a bank's relationship to a linked web site may help reduce legal risk to a bank arising from problems with services or products on the linked sites.
- Contingency planning
  Banks should have a contingency plan in the event of a disruption in its provision of electronic banking and electronic money services. The plan may address data recovery, alternative data-processing, capabilities, emergency staffing, and customer service support etc.

Many policies regarding managing e-banking risks have been also proposed by various practitioners and academics. For example, Stern (2001) outlines various policies on how to safeguarding customer information in e-banking.

By implementing appropriate security policies and measures, many risks related to e-banking can be reduced or eliminated. However, given the nature of e-banking risks faced by a bank, it is impossible to eliminate all the risks totally. We discuss below two important means of controlling e-banking risks.

Insuring e-banking risks

One way of mitigating the risks is to take insurance cover against eg. hackers, fraud, and network crashes etc.

As we mentioned earlier, problems that can plague on-line baking sites include denial-of-service attacks and website defacements, transmitting virus and privacy violations etc. These risks can never be eliminated regardless how stringent the risk policies might be. Each of these Internet-related banking risks has two-parts: the damage down to the bank itself through the loss of business or damage to reputation, and the damage done to customers or partners, for instance, if a company can't access its funds and loses a business opportunity, it may hold the bank responsible. Thus the traditional insurance policies may not adequately cover the risks that are associated with e-banking risks. For examples, O'Neill (2001) examined the traditional insurance policies a financial institution typically has. He claims that major coverage gaps exist in regard to e-commerce exposures. In order to filling the existing gaps, specialized coverage should be obtained and the coverage can include:

- Business income coverage that can replace not only the business income losses and additional expenses incurred as a result of interrupted services, but also can pay for the cost of investigating the reason for the loss of service.

- Impairment or interruption of service liability, which covers liabilities to third parties for e-business losses, including reasonable expenses incurred in the defense or appeal of claims.
- Intellectual property coverage to protect against the loss of proprietary information or software through deliberate or inadvertent misappropriation.
- Public relations coverage for the expenses incurred to help rebuild a financial institution's reputation following negative publicity resulting from e-business exposures.
- Electronic publishing liability to cover liabilities incurred from publishing information via the Internet. Covered losses include defamation of character, libel, and slander, copyright infringements, plagiarism and misappropriation of ideas.
- Rewards coverage, which pays for information leading to the arrest and conviction of any individuals committing or trying to commit any illegal act against the insured's e-business activities.

Insuring e-banking risks has not been used much in the past few years. It is not an overnight process and insurance policies are being quickly developed in this aspect. Without doubt, the overblown fears in e-commerce will benefit the insurance industry and more insurance policies for e-banking risks will be developed soon. For example, according to Anonymous (2001), the American Banks Associated sponsored insurance program announced a new product for financial institutions---the Internet Banking Package (IBPP). The IBPP underwritten by. Progressive, comprises an Internet banking liability policy and enhanced version of an existing financial institution bond. Optional endorsements include business interruption, public relations expense and cyber/network extortion coverage. The IBPP covers losses from acts such as: invasion of privacy; libel, slander and defamation or other actional verbal or written disparagement; loss or damage due to electronic data of a customer, denial of impairment or interruption of service; unauthorized access to a customer account; and infringement of copyright, misappropriation of ideas or plagiarism. The enhanced financial institution bond covers losses include: theft of electronic data or property by hackers; damage or destruction to electronic data or computer programs resulting from hackers; computer viruses and employee sabotage; fraudulent fund transfers initiated by fax, e-mail or Internet access (cf. www.aba.com).

It appears that the insurance for e-banking has just taken off in the US and how successful and effective the current offered polices remains to be seen. In other countries, such developments are yet due to appear. It appears that insuring e-banking risks will be an effective tool for managing e-banking risk in the future.

Hedging

The hedging concept in financial risk management is relevant to managing e-banking risks. In investments, hedging refers to buying an asset to reduce the risk in a portfolio. The term is common in futures and foreign exchange markets where traders use facilities available to protect themselves against future price or exchange rate variations. If someone bulk buys scotch whisky ahead of the budget in anticipation of a price rise in the budget, then he or she is hedging (provided the whisky is drunk—if it were bought to be sold, then the buyer is speculating). The benefit of hedging for an e-commerce firm is the reduction in e-commerce risks. That is, by hedging, an e-commerce firm can have a more stable income. For details regarding hedging e-commerce risks, which include e-banking risks, we refer to Li (20001).

MONITORING RISKS

Ongoing monitoring is an important aspect of any risk management process. For e-banking, monitoring is particularly important because the nature of the activities are likely to change rapidly as innovations occur, and because of the reliance of some products on the use of open networks such as Internet. Two important elements of monitoring are system testing and auditing.

Testing of systems operations can help detect unusual activity patterns and advert major system problems, disruptions, and attacks. Auditing (internal and external) provides an important independent

control mechanism for detecting deficiencies and minimizing risks in the provision of e-banking services.

Auditing (internal and external) provides an important independent mechanism for detecting deficiencies and minimizing risks in the provision of electronic banking. The role of an auditor is to ensure that appropriate standards, policies and procedures are developed, and that the bank consistently adheres to them.

SOME RISK MANAGEMENT PRINCIPLES

Risks do exist with e-banking as shown in the previous section. We have to live with them whether we like them or not. A sound risk management practice requires us to follow the three steps as illustrated above. Of course, risk management policies should be bank specific. The need of hedging or insuring will also vary from bank to bank. Nevertheless, there are some common characteristics with e-banking that call for some common management principles.

Since e-banking is based on technology that by its very nature is designed to expand the "virtual" geographical reach of banks and customers without necessarily require a similar "physical" expansion, market expansion can extend beyond national borders, which significantly increases cross-border cooperation challenges for bank supervisors.

Although the supervisory principles of traditional banking are applicable to e-banking, the amalgam of changes in technology and the degree of dependence exhibited by banks upon services providers and technological distributors mutate and magnify the typical levels of risk. This led to a report by the Electronic Banking Group of the Basel Committee on Banking Supervision which identifies 14 key risk management principles for e-banking (BIS 2001). Banking institutions and their supervisors should consider these principles when formulating risk management policies and processing the e-banking activities.

Board and Management Oversight

- *Effective management oversight of e-banking activities.* The board of directors and/or senior management should establish effective management oversight per the risks associated with e-banking activities, including the establishment of specific accountabilities, policies and controls to manage these risks. I addition e-banking risk management should be integrated within the institution's overall risk management process.
- *Establishment of a comprehensive security control process.* Banking institutions should establish authorization privileges, logical and physical access controls and adequate infrastructure security to maintain appropriate boundaries and restrictions on both internal and external user activities and properly safeguard the security of e-banking assets and information.
- *Comprehensive due diligence and management oversight process for outsourcing relationships and other third-party dependencies.* The board of directors and /or senior management should establish a comprehensive well-defined and on-going process for managing outsourced relationships and third party dependencies supporting e-banking including adequate due diligence that should be conducted before engaging into relationships with third parties.

Security control principles

- *Authentication of e-banking customers.* Appropriate measures and controls should be in place over authorization with e-banking systems, databases and applications.
- *Non-repudiation (accountability) for e-banking transactions.* Banking institutions should ensure non-reputation to hold users accountable for e-banking transactions and information.
- *Appropriate measure to ensure segregation of duties.* Appropriate measures are in place to ensure proper segregation of duties within e-banking systems, databases and applications.

- *Proper authorization measures and controls in e-banking system, databases, and applications.* Appropriate measures and controls should be in place over authorization within e-banking systems, databases and applications
- *Data integrity of e-banking transactions, records and information.* Banks should prevent unauthorized changes, ensure the reliability, accuracy and completeness of e-banking transactions, records and information.
- *Establishment of clear audit trails for e-banking transactions.* A clear audit trail should exist for all e-banking transactions.
- *Confidentiality of key bank information.* Banking institutions should take appropriate measures to ensure that key information remains private to themselves and is not viewed or used by those unauthorized to so.

Legal and reputational risk management

- *Appropriate disclosure for e-banking services.* To minimize legal and reputation risk associated with e-banking activities conducted both domestically and cross-border, banking institutions should provide their customers adequate disclosure with their websites, in order to assist them to make informed choice.
- *Privacy of customer information.* Banking institutions should take appropriate measures to preserve the confidentiality of customer information and ensure adherence to customer privacy requirements. Measures taken to preserve confidentiality and privacy should commensurate with the sensitivity of the information being transmitted.
- *Capacity, business continuity and contingency planning to ensure the availability of e-banking systems and services.* Banking institutions should implement effective capacity planning, business continuity, and contingency plans to ensure that e-banking systems and services are available to customers, internal users and outsource service providers when necessary.
- *Incident response planning.* Banking institutions should develop incident response plans to mange, contain and minimize problems arising from unexpected events including internal and external attacks that hamper the provision of e-banking systems and services.

Each of the above issues is discussed in detail in the report by the Basel Committee on Banking Supervision (2001). These principles provide a useful tool for managing e-banking risks, though they were not put forth as absolute requirements or even best practice. Each bank's risk profile is different and requires a tailored risk mitigation approach appropriate for the scale of the e-banking operations, the materiality of the risks present, and the willingness and ability of the institution to mange these risks. Thus the BIS report is by no-means a "one size fits all" approach to e-banking risk management.

CONCLUSIONS

As electronic banking becomes more widespread and complex, the need for banks to assess and manage e-banking risks will become ever more critical. The major risks in e-banking are addressed in this paper. In assessing the risk exposure of an e-banking business, we recommend a systematic approach by using a risk map. Risks such as Internet failure etc. can be effectively reduced by insurance or hedging, both are at early developing stage for the e-banking sector. Some financial products are needed for effectively managing the e-commerce risks. Finally, we address some common e-banking risk management principles.

REFERENCES

Anonymous (2001). Internet banking insurance, Risk management, October, pp10.

Anonymous (2002). Egg survey looks at the uptake of Internet Banking in Britain, Telecomworldwire, April 2.

Alstad J. (2002). Use your service edge to your online advantage, American Banker, New York, March 8.

Anonymous (2002). Egg survey looks at the uptake of Internet banking in Britain, Telcomworldwire, Coventry, April 2.

BIS (1998). Risk management for electronic banking and electronic money activities, Bank for International Settlement, March.

BIS (2001). Basel Committee on Banking Supervision, Risk Management principles for electronic banking, Bank for International Settlement.

BBC (2000). Safety fears for web banking. BBC News Tuesday, August 1[On-line].

Chance, D.M. (2001). *An introduction to derivatives and risk management,* 5th edition, Harcourt.

Claessens S., Glaessner T. and Klingebiel D. (2001). E-finance in emerging markets: is leapfrogging possible? Financial Sector Discussion Paper No.7, The World Bank.

Clark, P.B. (2001). Company hedges bet on ad buying, *B to B; Chicago*, April.

Creed, A. (2000). Western Union Site down after theft of credit card details, Newsbytes, Englewood, Colorado, USA, September 10.

Deutsche Bunddesbank (2000), Electronic banking from a prudential supervisory perspective, Deutsch Bundesbank Monthly Report, December.

De Young R. (2001). The financial progress of pure-play internet banks, BIS paper No 7.

Ferguson, K. (1999). Risk e-business, Forbes.com.

Fuhrman, A. (2002). Your e-banking future, Strategic finance, April, 24-29.

Greif, J. (2000). Risky e-business, *Association Management*, Nov.

Gutterman, S. (2000). Western Union web site is hacked, Associate Press, September 10.

Hutchinson, D. and Warren, M.J. (2001). A framework of security of authentication for Internet banking, 2nd International We-B conference, Perth.

Jablonowski, M. (2001). Thinking in numbers, *Risk Management*, New York, Feb.

Knight, W. (2000a). Barcklays security breach forces online service to close, ZDNet UK, Monday, July 31 [On-line].

Knight, W. (2000b). Barclays in security gaffe this week, ZDNet UK, Wednesday, August 2 [On-line].

Li, S. (2001). Some thoughts on managing e-commerce risks, Proceedings of the 2nd International We-B conference, Perth, Western Australia.

O'Neill, D. (2000). Evaluating banks' e-commerce risks, *American Agent &Broker,* November, 36-46.

Portter, M. (2002). Outsourcing Internet Security: A new business solution from America's community Bankers, Community Banker, January.

Robinson T (2000). Internet Banking: still not a perfect marriage, Informationweek.com, April 17.

Stern, M. (2001). Safeguarding customer information: The key to customer trust, ABA Bank compliance, November/December.

Sato, S., Hawkins, J. and Berentsen, A. (2001). E-finance: recent developments and policy implications, in Cohen, S. and J. Zysman (eds) *Tracking a transformation: E-commerce and the terms of competition in industries, Brooking Institution.*

Smithson, C.W. (1998). *Managing financial risk,* McGraw Hill.

Timmers, P. (1998). Business Models for electronic markets, *Focus theme*, vol. 8 No.2

Enhancing Airport Access Control Security

M. W. David

Cubic Corporation
Singapore
Email: mike.david@cubic.com

Abstract

This paper proposes the use of biometric features in contactless smart cards (CSC) to enhance airport access control security. It looks at biometric features like the user's fingerprint, face or iris to provide verification of the person presenting the card for access to a facility, or use of an application. The paper also suggests possible use of device features like uniquely identifiable semiconductor chips in combination with biometric features and personal data to support logical network integration with organizations outside the airport operations environment.

Keywords: Access Control, Airport Security, Biometrics, Contactless Smart Cards

INTRODUCTION

The events of September 11th 2001 have accelerated efforts to improve methods and means of individual identification (ID) for both physical and logical access. Air traffic controllers brought down every commercial plane in the air after the attacks in New York and Washington DC. If there had been a cyber attack at the same time that prevented them from doing that, the magnitude of the day's events could have been much greater [Gellman, 2002]. Anyone who has been to an airport since September 11th should be aware of the increased scrutiny given to passenger identification and baggage checks. Biometrics and smart cards have been proposed as tools to support this identification process. To date the United States has been lagging behind Europe and Asia in the introduction and use of smart card technology [Karlin, 2002]. However, the U.S. has been catching up, and initiated the Department of Defense (DOD) Common Access Card (CAC). Contactless smart cards are being used in transit systems in Washington DC and Chicago. DOD has created a Biometrics Management Office (BMO) to consolidate oversight and management of biometric technology for DOD [Woodward, 2001]. Questions linger about the security of smart cards and the reliability of fingerprint readers and facial recognition systems. However, developments in contactless smart cards, biometric products and semiconductor production methods are providing an effective means to use multiple features to support secure identification and verification [Vonverheid, 2002]. The paper proposes the integration of these technologies for use in multiple feature smart cards for identity authentication and access control, and suggest an immediate application for use in airport security systems.

THE CHANGING ENVIRONMENT

According to a October 2001 Rand report, the U.S. Supreme Court has found that a person does not have a reasonable expectation of privacy in those physical characteristics that are constantly exposed to the public, such as one's facial features, voice and handwriting. Secondly, current legal standards recognize that we are all subject to heightened scrutiny at our borders and ports of entry. This is based on "the longstanding right of the sovereign to protect itself by stopping and examining persons and property crossing into this country [Woodward et al., 2001]."

Outside the DOD, the Transportation Security Agency (TSA) is planning a multi-application smart ID card in North America as part of its Transportation Worker ID Card (TWIC) program. The TWIC will standardize on a single common ID card platform, containing a digital photo, hologram security layer, possibly a contact chip migrating to a contactless chip, a magnetic stripe, 2D barcode and a visible TWIC ID number. The biometric and other information to be contained on the card will be included in a central Department of Transportation (DOT) database and held there until background checks, drug tests and other security levels have been satisfied. There is a potential for issuance of 15 million TWIC [Vanderhoof, 2002].

SMART CARDS

Contact smart cards must be inserted into a card reader. They transfer data between the smart card and the reader/writer (r/w) unit through the use of six metallic connectors or contacts found on the surface of the card. When they make physical contact with the connectors to transfer data from the chip, the connectors receive an electrical voltage to power the MPU. The contact plate provides an input/output path for the transaction of data. However, they must be inserted into the r/w units that have movable parts, and these r/w units require extensive maintenance [Wilson, 2001].

The contactless smart card (CSC) has no surface contacts. The CSC has an integrated circuit (IC) chip and a radio frequency (RF) antenna embedded in it. The card must pass near an antenna to carry out a transaction. Power is transferred by an inductive loop using low-frequency radiation from an electromagnetic field, and an electrical-magnetic transformation occurs through the same antenna that transmits and receives data. The CSC gets its power from the RF field. The CSC uses proximity r/w units that have no moving parts, and are not as susceptible to maintenance failure under heat, humidity, dust or vibration. Typically, a CSC can process and transaction in 150 – 300 microseconds versus the 1.5 seconds for a contact card supporting a similar application [Wilson, 2001].

BIOMETRIC TECHNOLOGIES

Biometrics measures individuals' unique physical or behavioral characteristics to recognize or authenticate their identity. Common physical biometrics includes fingerprints; hand or palm geometry; and retina, iris or facial characteristics. Behavioral characters include signature and voice. Generally speaking, the less intrusive the biometric, the more readily it is accepted. However, certain users, religious groups and civil-liberties groups have rejected biometric technologies because of privacy concerns [Liu & Silverman, 2001].

Organizations should determine the level of security needed based on the application and the surrounding environment. This will influence which biometric(s) are most appropriate. That is, different biometrics may be appropriate for different applications, depending on perceived user profiles, the need to interface with other systems or databases, physical and environmental conditions and a host of other application-specific parameters. Table 1 provides a comparison of various biometric features against a number of characteristics. The comments in error incidence identify factors that can impact on the ability to properly accept or reject the biometric feature. The original table did not include "artificial" under fingerprints. This has been added based on research by T. Matsumoto of Yokohama National University, and is discussed in more detail below in the section on issues for biometric technologies.

Table 1: Comparison of biometrics

Characteristic	Fingerprints	Hand Geometry	Retina	Iris	Face	Signature	Voice
Ease of Use	High	High	Low	Medium	Medium	High	High
Error incidence	Dryness, dirt, Age, (**)	Hand injury, Age	Glasses	Poor Lighting	Lighting, age, glasses, hair	Changing Signatures	Noise, Colds, Weather
Accuracy	High	High	Very High	Very High	High	High	High
Cost	*	*	*	*	*	*	*
User acceptance	Medium	Medium	Medium	Medium	Medium	Medium	High
Required security level	High	Medium	High	Very High	Medium	Medium	Medium
Long-term stability	High	Medium	High	High	Medium	Medium	Medium

***The large number of factors involved makes a simple cost comparison impractical.**

Source. "A Practical Guide to Biometric Security Technology," IT Professional, Jan/Feb 2001 [Liu & Silverman, 2001]

** Authors' Comment: We propose adding, "artificial" based on "gummy finger" research: [Matsumoto et al., 2002].

BENEFITS IN A COMBINED SMART CARD/BIOMETRIC ID

Using biometric technologies with smart cards for ID system applications add a greater ability to identify people with minimal ambiguity. An ID system using a smart card, cryptographic functions and biometrics has significant security advantages.

The biometric template can be digitally signed and stored on the smart card at the time of enrollment and checked between the biometric capture device and the smart card each time the card is used.

The template and personal information on the card can be encrypted.

Cardholder authentication can be performed by the smart card comparing the live template with the template stored in the card. The biometric template never leaves the card, protecting the information from being accessed during transmission.

The smart card ID can challenge the biometric reader to ensure that a previously captured template is not being retransmitted in a form of playback attack

Smart cards have sufficient memory to store growing amounts of data including programs, one or more biometric templates, and multiple cryptographic keys to restrict data access and ensure that data is not modified, deleted or appended.

The smart card can be used to prove the digital identity of its cardholder using cryptographic keys and algorithms stored in the protected memory [SCA, 2002].

ISSUES FOR BIOMETRIC TECHNOLOGIES AND SMART CARDS

Researchers at Yokohama National University (YNU) have been working on artificial fingers for the past few years. The most recent results have reported attacks using artificial *gummy* fingers, namely artificial fingers made of cheap and readily available gelatin. These gummy fingers were accepted at extremely high rates by particular fingerprint devices with optical or capacitive sensors. The research revealed there are many possible attacks to deceive commercial fingerprint readers, even if the templates and communications are protected by secure measures. Most noticeably, eleven types of fingerprint systems accepted the *gummy* fingers in their enrollment procedures and also with the rather higher probability in their verification procedures [Matsumoto et al., 2002]. This should be a special concern for airport and other types of facilities that have large numbers or employees, and multiple access points, which may not be under visual observation by a guard or online surveillance system.

The Facial Recognition Vendors Test (FRVT) 2000 was sponsored by multiple U.S. government agencies to evaluate facial recognition systems. FRVT 2000 was developed using the evaluation methodology proposed in IEEE Computer article, "An Introduction to Evaluating Biometric Systems," by authors from the National Institute of Standards and Technology. The methodology proposed a three-step evaluation protocol: a top-level technology evaluation followed by a scenario evaluation, and finally an operational evaluation. FRVT 2000 performed a technology evaluation titled "Recognition Performance Test" and a limited scenario evaluation titled "Product Usability Test [Bone & Crumbacker, 2001].

The overall conclusion for recognition performance tests stated the FRVT 2000 showed that progress has been made in temporal changes, but developing algorithms that can handle temporal variations is still a necessary research area. In addition, developing algorithms that can compensate for pose variations, and illumination and distance changes were noted as other areas for future research. The FRVT 2002 is scheduled to begin in June 2002, and results announced sometime in the fall of 2002 [Blackburn et al., 2001].

Iris scanning is less intrusive than retina scanning. It utilizes a fairly conventional CCD camera element and requires no intimate contact between user and reader. As a technology it has attracted the attention of many integrators. It has been demonstrated to work with spectacles in place and with a variety of ethnic groups, and is one of the devices that can work well in the identification as well as authentication mode. The main practical problem facing deployment of iris scanning is getting the picture without being intrusive. Also, attacks could be made by a simple photograph of the target's iris, at least in unattended operations [Anderson, 2001].

Paul Kocher brought the threat posed to smart cards by power analysis to the attention of industry in 1998. He developed a specific signal-processing technique to extract the key bits used in a block cipher from a collection of power curves, without the knowing the implementation details of the card software. This technique has been called differential power analysis (DPA). Various defenses have been fielded, and new attacks have been mounted. This is an area of active research [Anderson, 2001].

DEALING WITH THE ISSUES

Although the results from Yokohama National University (YNU) provide ample room for improvement, they do not entirely eliminate fingerprint recognition as a valuable biometric feature for identifying an individual. Probably the most important lessons are the need to have more than one biometric feature, proper, live enrollment procedures and interface with the application system that will confirm the authenticity of the individual. That is, supervised and reliable enrollment supported by physical or alternative biometric authentication if necessary.

Figure 1 demonstrates the flow of how a live fingerprint is presented to the reader for capturing. The fingerprint feature data is then extracted by the sensors and related algorithms. The data is recorded on a smart card and/or database, and then compared again for correctness before the final result is registered and stored for future use. This registration should be done in the presence of a trusted party, and the registrant should have valid identification documents and clearances as appropriate to the situation. The risk of acceptance of a false fingerprint is greatly reduced when supervised systems make some check for liveness, however, supervision does not equal a biometric live and well test by high quality sensors.

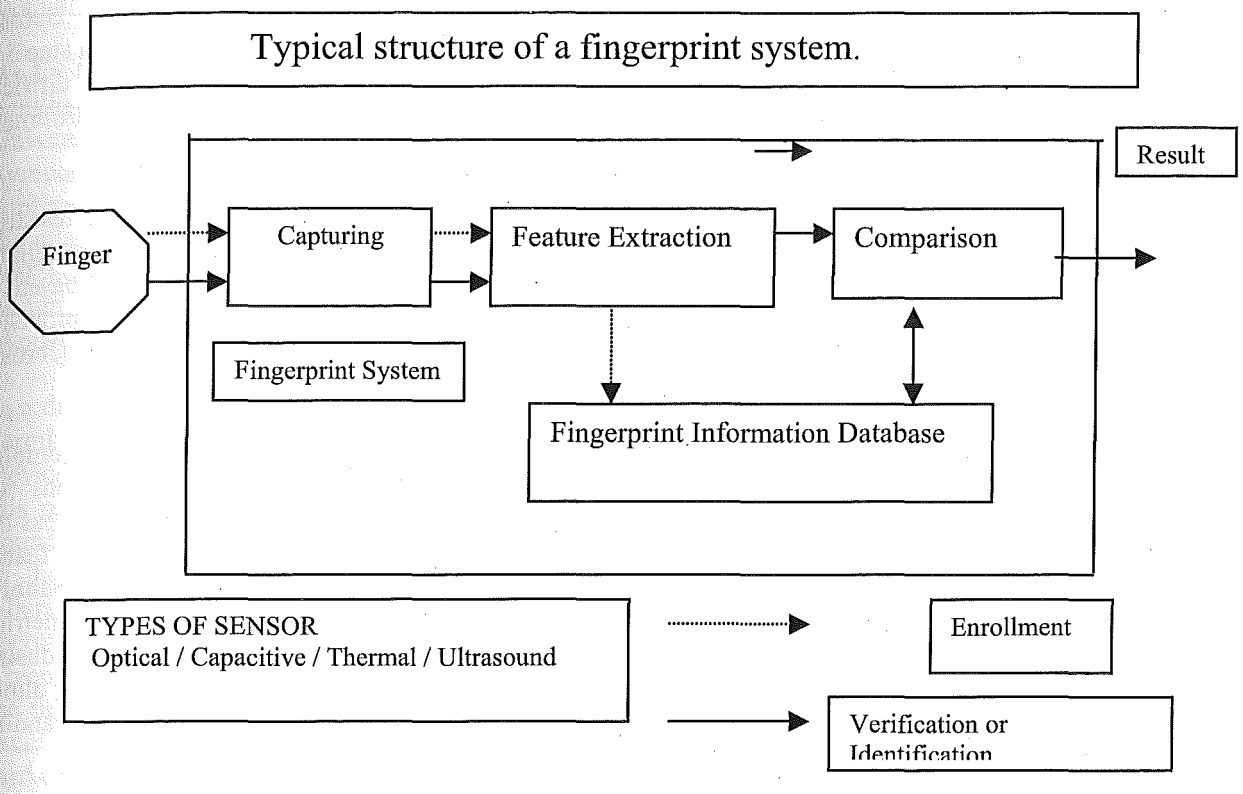


Figure 1: Sample Fingerprint Enrollment System. Matsumoto, T. "Case Study for User Identification" ITU-Workshop on Security, May 2002

Another example is the model proposed by the DOD Common Access Card (CAC) Biometrics Working Group below.

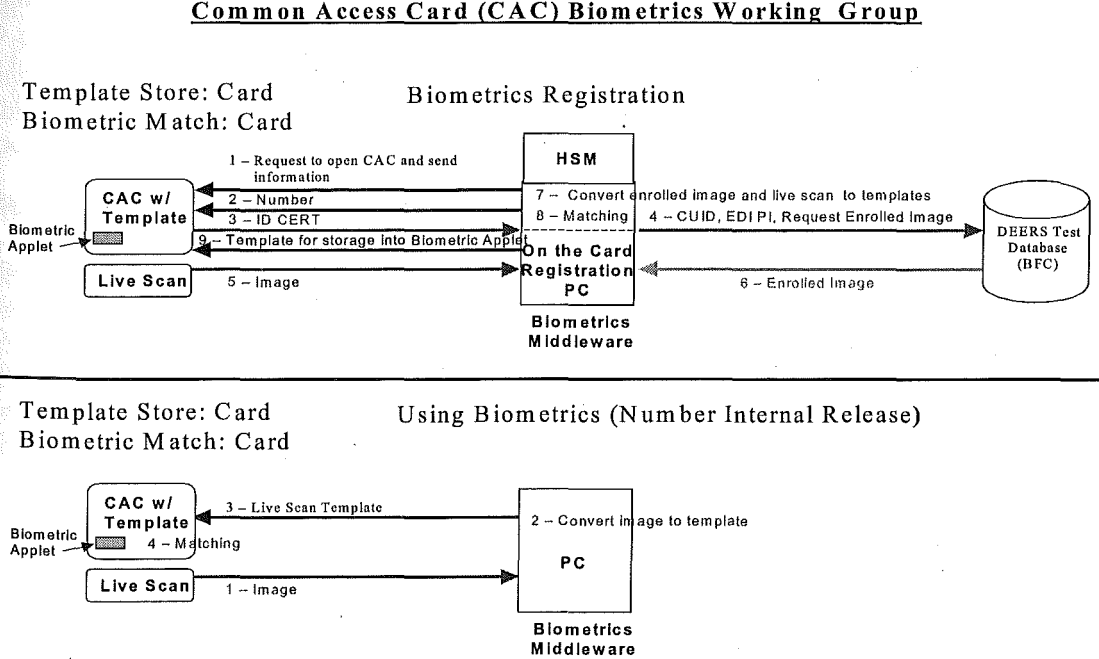


Figure 2: Scenario for biometrics registration from "Biometrics and the Common Access Card (CAC)" at CardTech / SecureTech, April 25, 2002 [Kocher, 2002].

Figure 2 depicts the flow of registration and use of the biometric template on a US Department of Defense (DOD) Common Access Card (CAC) smart card, and how it interfaces with the Defense Enrollment Eligibility Reporting System (DEERS). The key point is the uses of a live scan in step 5 of the registration process to avoid improper enrollment of a false identity or use of an artificial fingerprint. The enrollment processes in both Figures 1 and 2 emphasize the need for live enrollment, which should also be supervised by trusted personnel.

As noted in the FRVT 2000 report, facial recognition is an area that needs additional research. However, we believe improvements have been made since FRVT 2000, and anticipate more reliable algorithms and corresponding products will be available. One methodology could be based on Hausforff distance based models. Research in 2001 indicated this allows for an efficient approach to achieve fast, accurate face detection that is robust to changes in illumination and background [Jerosky et al., 2001].

There has also been encouraging research reported in early June 2002. One report on a complete scheme for face recognition based on salient feature extraction in challenging conditions was performed without any a priori or learned model. These features were used in a matching process that overcomes occlusion effects and facial expressions using dynamic space warping to align each feature in the query image, if possible, with its corresponding feature in the template or database [Sahbi & Boujenna, 2002]. Another report entitled "Understanding Ionic Image-Based Face Biometrics" describes a system for personal identity verification and recognition based on academic and industrial data sets. The experimental results reportedly show greatly improved performance reaching almost 100% recognition [Tisratelli et al., 2002]. Facial recognition is also less intrusive than iris scanning, and faster for use in high volume passenger areas like boarding gates.

A DOD Biometrics Fusion Center (BFC) product evaluation has looked at a number of iris scanning products. In general, these have been rated overall as "excellent" in their ability to meet DOD requirements [Kocher, 2002]. As far as is known, every human iris is measurably unique, even for identical twins. It is fairly easy to detect in a video picture, does not wear out, and is isolated from the external environment by the cornea. The iris pattern contains a large amount of randomness, and appears to have many times the degrees of freedom of a fingerprint. A possible solution to the impersonation problem is to design terminals that measure hippus, a natural fluctuation in the diameter of the pupil, which happens at about 0.5 hertz. Iris codes remain a very strong contender as they can, under the correct circumstances, provide much greater certainty than any other method that the individual in question is the same as the one who was initially registered on the system. They can meet the goal of automatic recognition with zero false acceptance [Anderson, 2001].

In practice, iris scanning is proving quite successful. The Amsterdam Schiphol airport has been using an Automatic Border Passage (ABP) system since October 2001. The security procedure for this system has two phases. The first is qualification and registration. This process includes a passport review, background check and iris scan that is encrypted and embedded on a smart card. The second phase identifies and verifies the registered traveler at the border passage checkpoint. The system reads the smart card and allows valid registered travelers to enter an isolated area. The traveler then looks into an iris scan camera so the iris can be matched with the data on the smart card. If the match is successful, the traveler exits, if it fails, the traveler is directed to the front of the standard queue for passport check [IBM, 2002].

The Canadian Customs and Revenue Agency will also begin to use iris scanners to speed air travelers through the country's busiest airports. Those who want the service will submit to a background security check, including a criminal record search. The International Air Transport Association (IATA) has indicated that scanning eyes is its preferred biometric choice. One important factor for the IATA is that using the eye as an individual's unique identifier appears to be the most socially neutral. For example, a Muslim woman could be identified without touching her or asking her to drop her veil [Akin, 2002].

The State Department has tried hand recognition and retinal scanning without success, but the technology is moving toward iris scanning and face recognition [Baker, 2002].

As for the smart cards attacks, for every attack, there is usually a suitable defense. It is the age-old paradigm of the defense versus the offense. For example, in May 2002, Ross Anderson of Cambridge University presented a paper at the 2002 Institute of Electrical and Electronics Engineers (IEEE) Symposium on Security and Privacy entitled *Optical Fault Induction Attacks* (OFIP). At the same conference, Simon Moore, also of Cambridge, presented a microchip design that could protect against the attack. Moore says: "No single point of failure will result in information being leaked" [Knight, 2002]. Anderson's own paper also offers a solution to the OFIP by using self-timed dual-rail circuit design techniques. This technology may also make power analysis attacks very much harder too [Anderson & Skorobogatov, 2002b].

In a different vein, Mitsubishi Electric Corp. has developed a semiconductor fabrication step that will permit every computer, smart card and semiconductor chip to have its own Artificial Fingerprint Device (AFD). Depositing a poly-silicon film on a large-scale integration wafer creates the AFD. In this process, crystals, called grains, form and are randomly distributed. This distribution of grain boundaries, which cannot be changed, is read by a thin film transistor (TFT) and a code is generated. In theory, 40 TFTs can provide one trillion numbers, and the TFT takes up very little space. Alteration and duplication are deemed to be impossible, and no additional cost is necessary. If it becomes standard for interactions between computers, smart cards and r/w devices, the host will have a record of the transaction and can identify whether the card and chip are the same as the one on which the fingerprint was registered [Vonverheid, 2002]. This means that even if data can be copied from a smart card, it will not be of any use for access unless the illegal user also has the original valid smart card to use with the r/w device. This should be especially helpful to network forensic analysis.

PROPOSAL DETAILS

The proposal is to use a contactless smart card (CSC) in conjunction with fingerprint; facial and iris biometric features supported by AFD as the next generation multiple purpose ID and access control card. The CSC is recommended due to the advantages of lower mechanical complexity of the r/w unit, thereby affording higher reliability and less maintenance in the field. The CSC has been proven in public transit applications worldwide in locations like Hong Kong, Pusan, Sao Paulo, Seoul, Singapore, Tokyo, and Washington DC. The high volume and passenger throughput requirements for public transit fit well with the needs of airport security and border control systems, where rapid, secure processing is needed.

Serious consideration should be given to the use of ferroelectric random access memory (FRAM) as the process and memory technology to be used in the next generation ID card. Currently, electronically erasable programmable read only memory (EEPROM) is the primary memory device used to store data in smart cards. However, FRAM is superior to EEPROM in many respects, including a write speed over 10,000 times faster. In addition, FRAM has lower voltage writes than EEPROM and FLASH; consumes about $1/400^{th}$ of the power EEPROM uses for writing data, and its rewrite endurance is 100,000 times greater [Ramtron, 2001]. A FRAM-based CSC system with 32 kilobytes of memory will soon be on the market, and could be used to support all of the proposed ID applications. The CSC includes an Internet Transaction System, which allows PC/laptop users secure logical access to online networks [Cubic Security, 2001].

The selective use of iris scanning and facial recognition should be used to support the growing needs of law enforcement, border security, transportation security, airline passengers and other potential users. The human fingerprint has its problems, but the overall pervasiveness of this feature in law enforcement, military, immigration and access control systems makes it difficult to ignore or dispense with. Therefore, there is a need to include fingerprint templates in the next generation ID card for compatibility with legacy systems.

The future application of the AFD on smart cards and semiconductor chips will provide additional ability to identify the individual, and verify the authenticity of the card. For example, an Internet Service Provider (ISP) or wireless telephone company could use AFD to identify its legal subscribers and provide service only to them. In support of computer forensics, the AFD would become the iris or fingerprint of the device, and provide investigators with the ability to determine the origins of an incident. A near term application for this system could be in support of airport and airline security systems being planned in the US and around the world. A generic outline of one is provided in Figure 3.

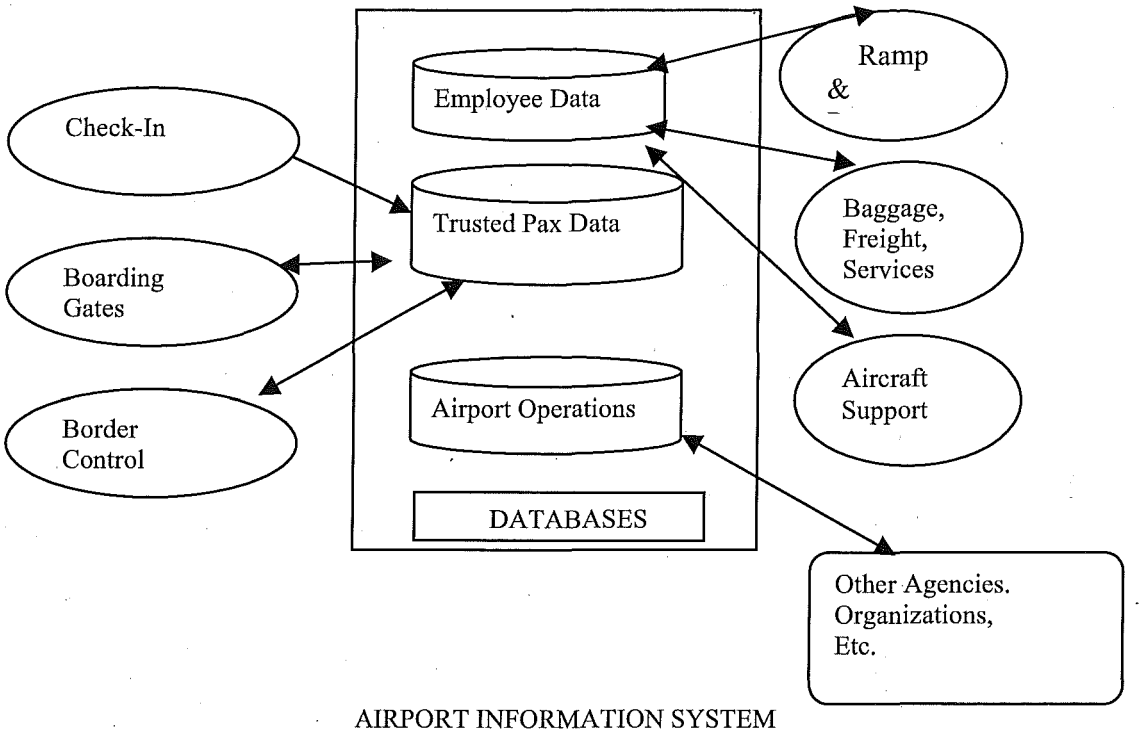


Figure 3: Proposed outline for an airport identity and access control information system.

A key factor for a secure access control system is, first and foremost, proper and accurate enrolment procedures as suggested in Figures 1 & 2.

The system outlined in Figure 3 above should be designed and implemented not just for the airport and airlines, but also with due consideration for related organizations like immigrations, customs, law enforcement and intelligence agencies. The key elements of the airport information system (AIS) in Figure 3 are the employee and trusted passenger (Pax) databases. The trusted pax database would interface with the locations where passengers enter and exit the system. That is, check-in counters or kiosks, boarding gates and border control points. The employee database would support access control for the full time and contractor personnel who work at, or perform services in support of the airport and airlines. When authorized, airport employee cards could also be used to support secure logical access to various airport information systems like reservations, check-in, air traffic control and other key computer systems. Integrating or interfacing the AIS databases with outside organizations would ensure that the most recent terrorist, criminal and security alert information is available to support passenger and employee checks, and determining the level of security in effect at the airport.

The use of multiple biometric features would allow the selection of devices and applications to fit the security, economic and social needs of the specific subsystem and its environment. For example, counter check-in may only require a fingerprint authentication, since the attendant will do the facial

check against the photo ID. Kiosk/automated check-in and gates should require at least fingerprint and facial authentication. However, immigration, air traffic control, aircraft maintenance area access verification may be better served by the higher level of iris scanning security. Therefore, the CSC and its related reader/writer system will need to be able to download various software biometric algorithms to support multiple levels of security and access control requirements. Table 2 below suggests some of the possible combinations of biometric features with the CSC ID.

Table 2: Proposed Combinations of Biometric Features with Contactless Smart Card ID

	Fingerprint	Facial	Iris
Check-in Counters with Attendants	X	O	
Check-in kiosk / terminals (un-attended)	X	X	
Boarding or other automatic gates	X	X	
Ramp & Runway Areas	X	X	O
Baggage & Freight Areas	X	X	O
Aircraft Maintenance, Support, Services	X		X
Border Control	X		X
Logical Network Access	X		X
Access & Coordination with Outside Agencies	X		X

Note: X – Highly Recommended O – Suggested Option

There are advocates for the use of voice recognition, because people are accustomed to speaking into phones, and there is no intrusive process or physical contact required. However, our proposal is focused on an airport system, and tied to a multiple feature smart card. In the first case, airports are very noisy places. The surroundings include numerous types of public address system announcements, passenger and staff activities and conversations; vehicle sounds and of course jet engines. Voice recognition also often requires some form of online interaction, which would necessitate a large database to match against. This process would take multiple seconds to perform, and might have to be repeated. Contactless smart cards can conduct on-card template authentication in a matter of milliseconds. This rapid processing speed is essential to support passenger throughout, staff efficiency and responsiveness.

Secure logical access is important to the prevention of illegal intrusions into databases and supervisory control and data acquisition (SCADA) systems. Access to command of such systems could permit terrorists to control or interfere with the air traffic control systems. Iris scanning, using mini charged couple device (CCD) cameras would be an excellent supplement to the existing logical access systems offered by the use of passwords and fingerprint readers.

Table 2 uses fingerprints instead of hand geometry because of the widespread availability of affordable, compact readers that support not only physical, but also logical network access. Retina scanning and signatures are not included because they are difficult to use in an airport where there is a high level of physical activity, and need for rapid movement of large numbers of personnel.

SUMMARY

The paper has noted the post September 11th 2001 environment in the US is calling for increased security at airports. This has created a search for improved identification and access control measures. The paper has looked at the smart card and biometric technologies that may be able to meet these needs. None of the solutions are perfect, and the paper has identified issues related to the use of smart cards and biometrics. Figures 1 and 2 offer some solutions to the threat of artificial fingerprints. Figure 3 outlines a generic airport identity and access control information system to improve airport security. Table 2 proposes how to combine biometric features with contactless smart cards to

implement the improved access control system outlined in Figure 3. The paper does not advocate eliminating human interface and contact with passengers. There should still be guards, inspectors and airline employees in the system that can scrutinize and assess suspicious or unusual behavior. However, it does propose technical methods to support and improve physical and logical access security.

FUTURE RESEARCH

There is a need for additional research on each of the subsystems in the generic outline presented in Figure 3 above. These subsystems must be capable of fully integrated operations at the database level to provide an overall contribution to the larger requirements of Homeland Security and counter terrorist operations. This leads to the need for more work on interoperability, cooperation and coordination, all of which are suitable topics for continuing research and implementation.

References

Akin, D. (2002): Customs Set to Use Iris Scans at Airports, http://www.iridiantech.com/news.php, Feb 2002

Anderson, R.J. (2001): Security Engineering-A Guide to Building Dependable Distributed Systems', Wiley 2001

Anderson, R.J. & Skorobogatov, S (2002): Optical Fault Induction Attacks, http://www.ftp.cl.cam.ac.uk/ftp/users.rja14/faultpap3.pdf, May 2002

Baker, D.A.(2002): Pentagon Endorses Biometrics to Enhance Computer Security, http://nationaldefense.ndia.org/article.cfm?Id=521, May 2002

Blackburn, D., Bone, M., Phillips, P.J. (2001): Facial Recognition Vendor Test 2000, http://www.frvt.org/DLs/FRVT_2000.pdf, Feb 2001

Bone, M.& Crumbacker, C. (2001): Facial Recognition – Assessing Its Viability in the Corrections Environment, http://frvt.org/DLs/FRVT2.pdf, Jul 2001

Constantinou, M. (2002): Identity Politics, http://www.sfgate.com/cgi-bin/article.cgi?file=/chroncile/archive/2002/03/03CM153735.DTL, Mar 2002

Cubic Security (2001): http://www.cubic.com/CSD/go-card.htm, Dec 2001

Gellman, B. (2002): http://www.washingtonpost.com/ac2/wp-dyn/A50765-2002Jun26?lanuage=printer, 26 Jun 2002

IBM Nederland (2002): http://www.ibm.com.news/nl/20022604_nl_nl_ibm_looks_airline_security_in_the_eye.html, Apr 2002

Jesorsky, O., Frischolz, R.W., Kirchberg, and K.J. (2001): Robust Face Detection Using the Hausdorff Distance. In: Proc. Third International Conference on Audio and Video-based Biometric Person Authentication. Lecture Notes in Computer Science, LNCS-2091. Springer-Verlag, Berlin Heidelberg New York (2001) 90-95

Karlin, S. (2002): Smart Cards, Widely Used in Europe, Migrate to United States, http://www.spectrum.ieee.org/spectrum/mar/02/departments/ncard.html, Mar 2002

Knight, W. (2002): Camera Flash Opens Up Smart Cards, http://newscientist.com/news/print.jsp?id=ns99992273, May 2002

Kocher, R.W. (2002): Biometrics and the Common Access Card (CAC), http://c3i.osd.mil/biometrics/, April 2002

Liu, S. & Silverman, M. (2001).: A Practical Guide to Biometric Security Technology, http://computer.org/itpro/homepage/Jan_Feb01/security3.htm, Feb 2001

Matsumoto, T. Matsumoto, H., Hoshino, S., Yamada, K. (2002): Impact of Artificial "Gummy" Fingers on Fingerprint Systems, http://www.spie.org/Conferences/Programs/02/pw/conds/4677.html, 2002

Ramtron (2001): Ramtron and Fujitsu to Jointly Develop Advanced 0.35 Micron FRAM Memory Process, http://www.ramtron/com/news/, 2001

Sahbi, H. & Boujenaa, N. (2002): Robust Face Recognition Using Dynamic Space Warping. Lecture Notes In Computer Science, LNCS-2359. Springer-Verlag, Berlin Heidelberg New York (2002) 121

Smart Card Alliance (SCA, 2002): Smart Cards and Biometrics in Privacy-Sensitive Secure Personal Identification System, http://www.smartcardalliance.org, May 2002

Tisratelli, M., Lagorio, A., Grosso, E. (2002): Understanding Iconic Image-Based Face Biometrics. Lecture Notes in Computer Science, LNCS-2359. Springer-Verlag, Berlin Heidelberg New York (2002) 19

Vanderhoof, R. (2002): President's Letter May 2002, http://www.smartcardalliance.org/newsletter.may_2002.htm

Vonverheid, E. (2002): Artificial Fingerprints-on-a-Chip Could Discourage Computer Crime, http://www.caffeine.ieee.org?INST/mar02/ffinger.html, Mar 2002

Wilson, Chuck. (2001): Get Smart: The Emergence of Smart Cards in the United States and Their Pivotal Role in Internet Commerce. Mullaney Publishing Group, 2001, Page 89-90

Woodward, J. D., Webb, K.W., Newton, E.M., Bradley, M., Rubenson, D. (2001): Army Biometrics Applications: Identifying and Addressing Sociocultural Concerns, http://www.rand.org/publications/MR/MR1237, 2001

Woodward, J.D. (2001): Biometrics – Facing Up to Terrorism, http://www.rand/org/organization/ard/, Oct 2001

Insights on the Development of a Robust C2 Information Infrastructure from Complex Adaptive Systems

Jill Slay

Advanced Computing Research Centre
School of Computer and Information Science
University of South Australia, South Australia
Email: Jill.Slay@unisa.edu.au

ABSTRACT

Modern C2 systems are difficult to secure because they are so widely distributed, rely on human agents from a wide range of national, service and professional cultural backgrounds and are built on non-compatible platforms with a poorly integrated software and information architecture. This paper draws on insights from Complex Adaptive Systems (CAS) in modelling of military Command and Control (C2) systems so to develop systems that are both organisationally and technically secure.

Keywords: Information warfare, C2, Complex Adaptive Systems.

INTRODUCTION

When military practice over the last twenty years is examined, it can be found that a Revolution in Military Affairs (RMA) has occurred with a move towards heavy reliance on distributed information systems in a military context. Those who carry out Command and Control (C2) activities in this IT rich environment now draw on various paradigms, including information theory and systems engineering, for their understanding, in advancing their body of knowledge and for the successful completion of military operations. Those who carry out research and practice in the C2 area draw on the 'art' of the discipline, as the term suggests, to develop their skills in commanding and controlling their military forces. Good C2 depends on many factors including operational culture and individual leadership styles. Issues in C2 have become closely associated with Information Warfare (IW) and Cyber Attack since (Warren & Hutchinson, 2001) 'in some respects IW is a subset of the RMA which is also concerned with the military application of new technologies'. In a wider sense though, IW is the action of attacking (and defending) a nation's information infrastructure (Schwartau, 1994) including its government networks, financial and telecommunications infrastructure as well as its military systems. (Warren & Hutchinson, 2001).

As has been stated elsewhere (Slay, 2002), since September 11th 2001 Western governments, their defence forces and defence researchers have been challenged as they have been forced to take seriously attack scenarios, both physical and cyber terrorist in nature, which originally would have been discounted as far fetched or science fiction. Recent work (Vatis, 2001)) that postdates the inception of the War on Terrorism also predicts an increase in cyber attacks on both military and non-military systems, accompanying physical kinetic warfare during the War on Terrorism, and urges Governments to follow 'best practices' in the defence of their information infrastructures.

These kind of best practices include, as might be expected, application of strict policies and the use of high-fidelity intrusion detection systems and firewalls. It has, however, been pointed out (Quirchmayr & Slay, 2001) that while effective security measures continue to be viewed as an add-on to a system and not as an essential feature of the initial design, an information system will still be at risk. Isolated efforts to fix security problems in a random fashion clearly are not the ideal solution and so security

issues must be addressed at a higher level, the design level. Before the security issue can be addressed at the design level there must be some reasonably sure method of modelling the system.

Military information systems have tended to develop rapidly and very often each service is using multiple hardware and software platforms that provide large-scale integration problems. The defence purchasing cycle is an incredibly long one (a recent purchase of helicopters by the Australian Government has taken more than twelve years between initial design and delivery) and there is an enormous lag between the requirements and the installation phase of the software and hardware development cycle. System components need to be used in times of peace, but in war response times need to be much quicker and a crisis situation can be reached very rapidly. Elements of the systems need to be highly mobile, extremely secure and thoroughly integrated. This situation is also made more complicated by the move towards joint and coalition operations that raise the need to deal with complex organizational and national cultural barriers as well as the technical ones of incompatible software and hardware raised above.

A major issue that has been, and continues to be, faced within the modern military context is how a historically hierarchical form of leadership and management, which displays a heavy reliance heavily on network-based information systems, can operate in an environment where technical skills and expertise now reside much lower down the chain-of-command. It is now common to hear discussion of Network-Centric Warfare that has transformed military operations and military decision-making in a radical manner.

The outcome then is that it is very difficult to secure modern C2 systems because they are so widely distributed, rely on human agents from a wide range of national, service and professional cultural backgrounds and are built on non-compatible platforms with a poorly integrated software and information architecture. An important task is thus to understand how to model these systems so as to identify where, and for what reason, they are insecure and then to develop, from an organisational and a technical perspective, secure systems.

METHODS FOR MODELLING DISTRIBUTED INFORMATION SYSTEMS

As has been shown previously (Slay, 2001), the use of soft systems methodologies allow us to identify internet-mediated distributed systems as complex socio-technical systems since they can be idealised as open systems which depend on the technology, the sentiments of the members, and the organisational environment (Checkland, 1981). Although the system is organised to focus on a primary task (military C2) this cannot be separated from the environment and the social factors, including cultural ones.

In previous work (Slay, 2001) helpful perspectives from Checkland have provided a reasonable method for gaining an understanding of such a system by producing an overview of the system from several perspectives. This allowed subjective and objective impressions of the system to be incorporated into a bricolage, a rich picture, which allowed the inclusion of human agents, the problems, conflicts and other seemingly 'soft' aspects of the system so as to determine the areas, which need improvement. Formalisation of these conceptual models allowed for solution of the problems within the system.

In this same work (Slay, 2001) issues of complexity are dealt with by the referring to Kline (1995) whose work gives an understanding of how credible perspectives on any specific complex system under examination may be derived. He maintains that three foundational perspectives of a multifaceted and hierarchical system are a synoptic, a piecewise and a structural perspective.

These views include:

- A **structural view** is one that provides details of how a particular system fits together within its hierarchy and provides information on the relationship between local and global effects within the system.
- A **piecewise view** is one that looks at the smallest portions of a system that might be relevant in providing information to aid in the solution of any particular problem.
- A **synoptic** view is an overview, which extracts, synthesises and thus maps a desired property of the system.

Other work (Slay (2002)) in modeling the military command and control enterprise uses the Generalised Enterprise Reference Framework and Architecture (GERAM) (IFIP-IPAC, 1999) to deal with understanding the complexity of a military command and control system. This 'defines a tool-kit of concepts for designing and maintaining enterprises for their life history'. As such it is not itself an enterprise reference architecture but a method of organising previously existent architectures and integrating current knowledge. The GERAM framework allows the modelling and visualisation of enterprises and can be applied to a military enterprise in general, and command and control in particular. Its power is in its provision of a standard enterprise framework that is recognised, and thus can be applied, internationally and in its modelling views that allow the identification of human and automated activities and differentiates between software and hardware tasks.

The weakness of the above three analyses (soft systems methodologies, Kline's three orthogonal perspectives on a complex system or an enterprise reference architecture) are that they fail to recognise that a military information system, a C2 system, probably more than any system, is an adaptive complex system. These analyses can only provide 'snapshots' in time (Lycett, 1999) which do not necessarily provide any misinformation but, because they are static, do not provide helpful insights into why military information system fail and, potentially, why wars are lost.

It has been pointed out (Schmitt (1999)) that military theorists have routinely turned to science to seek 'analogies and explanations' for the solution of battlefield scenarios and various defence researchers have pointed to the concept of complexity for understanding war, and military response systems. Schmitt comments that 'Newtonian science' no longer has an explanation to modern warriors but war (and therefore C2 and, by inference, modern military information systems) is an 'hierarchy of complex systems nested inside one another' with all the typical evolutionary properties such as emergence, uncertainty, adaptivity and flexibility.

It was also stated by American Defence C2 researchers and analysts (CCRP, 2001) that in the attack on the World Trade Centre 'disaster and crisis environments created difficult conditions for human decision makers and the socio-technical systems they interact with ... the design of effective response systems requires a socio-technical approach'. They believe that the solution to what they identify as their own complete organisational failure is a 'new pattern of interaction ...a response system (which)...can be seen to evolve through self organizing behavior of component units into a complex adaptive system'. Different techniques must then be used to model, develop and design such a response system.

FOUNDATIONS IN MODELLING COMPLEX ADAPTIVE SYSTEMS

Complex adaptive systems (CAS) are dynamic systems able to adapt and change within, or as part of, a changing environment. The work of Holland (1995) is foundational in the understanding of Complex Adaptive Systems (CAS). His basic treatise is that 'our intuition' (p.5) is that the world consists of complex patterns of behaviour whether one examines biology, epidemiology, economics, ecosystems or, as here, complex information systems. His work, and that of the Santa Fe Institute, has identified that many of these very different kinds of CAS share the same kinds of 'features and puzzles'. His search has been to extract general principles and apply them to various CAS with the

linking factor being that each of these systems exhibits 'coherence under change'. He identifies as a general feature of many CAS that they have 'lever' points where a small input can have 'major predictable directed change' (p.5). He also notes CAS 'abound in nonlinearities' that makes theory difficult to extract and so it is only in a cross-disciplinary comparison of CAS that makes possible the extraction of common properties.

Holland (Holland (1995)) explains that every CAS consists of active agents that are divergent in form and behaviour. Agents' behaviour is controlled by a series of stimulus-response rules and, once a set of stimuli and responses have been determined, then the kinds of rules a particular agent can have also been developed. These agents adapt by changing their rules as experience accumulates. He proposes that every CAS has seven basics. These are four properties and three mechanisms that are common to all CAS. The properties briefly described are:

Aggregation
A CAS can be analysed and similar parts grouped together and treated as equivalent. The aggregates that are formed can in turn act as higher-level agents – *meta agents*

Nonlinearity
A CAS is a non-linear system and non-linear interactions always make the behaviour of the aggregate more complicated than would be predicted by summing and averaging the behaviours of the aggregated agents

Flows
Flows in a network vary over time and have a multiplying effect. Flows are cyclical and recycling can increase output

Diversity
A CAS exhibits complex evolving patterns of diversity caused by progressive adaptation within the system

Mechanisms developed to deal with the properties are:

Tagging
CAS use tags to manipulate symmetries and facilitate selective interaction

Internal Models
Tacit internal models prescribe a current action under the anticipation of some desired future state. Overt internal models are used as the basis for explicit internal examination of alternatives – *lookahead*

Building Blocks
Building blocks – a limited range of reusable components- are used to generate internal models of CAS.

From these basic mechanisms and properties Holland (Holland, 1995, p.95) developed a generic framework for the examination of CAS and developed ECHO software for computer-based modeling and simple simulations of CAS.

RATIONALE FOR APPLYING CAS IN A MILITARY CONTEXT

Some researchers from a wide range of disciplines use computer based modelling to quantify effects within CAS as Holland did, while others (Axelrod & Cohen, 2000) use CAS to understand organisational and other kinds of social systems. In their analysis there are three key processes in a CAS: Variation, Interaction, and Selection. They see these as 'interlocking sets of concepts that can generate productive actions in a world that cannot be fully understood'.

While Axelrod & Cohen accept that 'patterns one sees in biology are not always found in other Complex Adaptive Systems' they emphasise the contextual forces determining interaction patterns. They see the CAS approach is a method of developing a perspective on the world that provides a set of concepts, a set of questions, and a set of design issues.

This research approach is also taken by Mitleton-Kelly (2002, p1)) and the London School of Economics Complexity and Organisational Learning Research Programme. It differs inasmuch as her work takes *'generic characteristics of complex adaptive systems'* and relates them to social systems and organisations.

APPLY COMPLEXITY THEORIES TO MILITARY INFORMATION SYSTEMS

It has been shown above that Military Information systems are complex systems. For reasons of interoperability between different services (nationally) and between different countries (in coalition operations) they are always in transition, evolving and being transformed for the specific needs of new operations.

In every CAS there are agents that interact with each other. Here we can identify Human Agents and Electronic Agents. Agents are placed in sites and cannot move within this local territory that makes up their world. All interactions and the structure of the agents are centred around a renewable resource – in this case the resource would be different categories of secure and insecure information. Agents can interact (trade, mate, fight)– in this case exchange information.

Military systems consist of the following features (Moore, 2001) expressed, where possible, in the language of CAS:

- Environment: a defended network
- Nodes: Critical nodes in the defended network
- Cyber security measures: firewalls, guards, encryption, authentication
- Cyber operating conditions: bandwidth, Internet activity.
- System constraints: static, outages, corrupt data
- Communication pathways: satellite communications, fiber, common carrier, networks, routers, switches, civil communications infrastructure
- Human Agents: Commanders, staff, users, civilians
- Electronic Agents: Command Centres, computers, databases,
- Weapons: Exploits, viruses, Trojans, malicious code, sniffers, system tools
- Targets: key network nodes, links, databases

Features of the model also include:

Replication: In a generic CAS agents replicate if they store enough resources – it is difficult to see how this applies to human agents but, in terms of electronic one, this would mean that there would be need for more databases and electronic storage as more secure information was accessed

Selective Interaction: Agents can differentiate between other agents before interacting – this means human agents could differentiate between friendly and enemy human or electronic agents.

Resource transformation: different types of resource can be transformed (here between one category of information and another).

CONCLUSION - IMPLICATIONS FOR MILITARY SYSTEM SECURITY

Qualitative data on these complex socio-technical systems can help identify organisational and socio-cultural issues leading to threat. This is similar to the work of Mitleton-Kelly (referred to above). In her approach to complex evolving and highly technological systems, Mitleton-Kelly worked with several complex British engineering enterprises such as Rolls Royce, to enable them to manage the 'co-evolution between changing business processes and information systems development'. She reports successful tests of aiding organisations, knowingly applying complexity theory within problematic situations, to co-evolve within a changing environment.

The quantification of this kind of qualitative data can lead to the development of rules that allow the prediction of the type of behaviour pattern which is being dealt with by a system (see Quirchmayr, G & Slay, J., (2001)). This work examined users communication patterns through the examination of web-server log. This is formally based on Rule Extraction Using Rough Sets When Membership Values are Intervals (de Korvin et al., 1998). This kind of algorithmic modelling of threats to the system can thus be used to generate rules for software agents which can be used in a proactive or reactive way to secure distributed systems (Busuttil & Warren, 2001).

FURTHER WORK

Once it can be established that War and Information Warfare are complex adaptive systems then a vast cross-disciplinary body of work in CAS, produced over the last twenty years, can be drawn upon to provide insight and analysis which will aid in the development of a robust and secure government, commercial and military IT infrastructure. This paper has only identified one or two very specific areas of research in CAS that offer insights that can be applied in the modelling and development of CAS. Further work will examine other cross-disciplinary applications of CAS that may prove to be more appropriate in this context.

REFERENCES

Axelrod, R. and Cohen, M.D (2000). Harnessing Complexity: Organizational Implications of a Scientific Frontier. Simon and Schuster, New York.

Command and Control Research Program Office of the Assistant Secretary of Defensefor Command, Control, Communications and Intelligence (CCRP). (2001). Sensemaking Symposium Final Report. Evidence Based Research Inc. (Online: accessed 26/02/02). URL: http://www.dodccrp.org/

Checkland, P. (1981). Systems Thinking, Systems Practice. Wiley, Chichester, UK.

de Korvin, A., Quirchmayr, G., Hashemi,S. & Kleyle, R. Rule Extraction Using Rough Sets when Membership Values are Intervals. Proc. of the 9th International Confernce on Database and Expert Systems Applications (DEXA 98), Vienna, Austria, August, 1998.

Holland, J.H. (1995). Hidden Order: How Adaptation Builds Complexity. Perseus Books, Cambridge, Mass.

IFIP-IFAC Task Force. (1999). Generalised Enterprise Reference Architecture and Methodology. IFIP-IFAC Task Force.

Kline, S.J. (1995). Conceptual Foundations for Multidisciplinary Thinking. Stanford University Press, Stanford, CA, USA.

Lycett, M., Kanellis, P. and Paul, R.J. (1997). Philosophical Directions Of Information Systems Development. Proceedings of the Association for Information Systems, 1997 Americas Conference.

Mitleton-Kelly, E. (2002). The Principles of Complexity and Enabling Infrastructures. In Complex systems and Evolutionary Perspectives of Organisations. (E. Mitleton-Kelly, ed.) Elsevier. (Online: accessed 26/02/02). URL: http://is.lse.ac.uk/complexity/EMK_The_Principles_of_Complexity.pdf

Mitleton-Kelly, E. (1997). Organisations as Co-evolving Complex Adaptive Systems. British Academy of Management Conference, 1997.

Moore, R.A., Williams, J.K and McCain, C. (2001). Intelligence Preparation Of The Information Battlespace- A Methodical Approach To Cyber Defense Planning. Proceedings of the 2001 IEEE Workshop on Information Assurance and Security, West Point, NY, 5-6 June 2001.

Quirchmayr, G & Slay, J.,(2001) A BPR-Based Architecture for Changing Corporate Approaches to Security. Australian Security Research Symposium. Perth, Australia. July 2001

Schmitt, J. (1999) Command And (Out Of) Control: The Military Implications Of Complexity Theory. National Defence University. (Online: accessed 26/02/02). URL: http://www.ndu.edu/inss/books/complexity/ch09.html

Schwartau, W. (1994). Information Warfare: Chaos on the Information Superhighway. Thunder's Mouth Press, New York.

Slay, J. (2001). Culture And Sensemaking In Information Warfare. 2nd Australian Information Warfare & Security Conference 2001, Perth.

Slay, J. (2002). A Cultural Framework for the Interoperability of C2 systems. European Conference on Information Warfare and Security. London, July 8-9 2002

Warren, M.J. & Hutchinson, W. (2001) Information Warfare: Fact or Fiction? 2nd Australian Information Warfare & Security Conference 2001, Perth.

A suggestion for a holistic (descriptive) approach to modelling physical security decisions

Z. Alach,[1] and C. L. Smith[2]

Security and Applied Technology Research (SATR) Group
School of Engineering and Mathematics
Edith Cowan University
Joondalup, Western Australia
[1]Email: zachary.alach@health.wa.gov.au
[2]Email: clifton.smith@ecu.edu.au

**Machine Vision research Group*
Department of Electrical and Electronic Engineering
Nottingham Trent University
Nottingham, England

ABSTRACT

A physical security decision model may provide the necessary supporting documentation to justify expenditure on physical security assets. This paper presents an overview of how some theoretical methods from other disciplines (Phenomenological Sociology, the Delphi technique, and Structural Equation Modelling) could be applied to decision modelling of physical security requirements. If a method of inquiry utilising these techniques/approaches, offers an alternative solution in identifying the universal elements of physical security and their associated relationships, then policy decisions may be automated based on relative values of any subset of these elements. A common theme through the literature of a universal (user) language, could effectively be used as the descriptive text within a universal element database (as defined by expert opinion) to link the model data to a physical environment. The determination of a universal element database, the relative importance of each element and any causal relationships associated with two or more elements may be the starting point for the design of a holistic physical security model.

Key words: *Phenomenological sociology, physical security model, modelling, decision theory, Delphi technique, SEM, lisrel, security technology.*

INTRODUCTION

The concept of Defence in Depth is a fundamental principle in the field of physical security. Defence in Depth is created with a series of rings or barriers (including psychological, technological, physical and procedural barriers) that may surround an asset or facility and typically associate the terms; deterrence, detection, delay and response as required components (Smith & Robinson, 1999, pp31). Achieving an adequate balance of components in a physical security system may be supported along with any decisions documented by an appropriate computer model.

This paper will approach some of the difficulties encountered in the task of modelling decisions for physical security requirements and will attempt to incorporate theory that may offer an alternate approach to current treatments. A suitable physical security model can be used by both decision makers and security consultants to evaluate existing structures or proposed changes to current security

environments. The benefit of such a model may be the output of a consistent and fully documented decision structure, using an input of appropriately gathered data.

This paper will discuss some theories of decision including the Delphi technique, present a method of inquiry using Phenomenological Sociology that addresses the universal elements of physical security, and incorporate Structural Equation Modelling to analyse data for decision making in physical security.

METHODOLOGY

The methodology of this paper is aimed at suggesting an extension of theory for physical security decision modelling. To build a model that can attempt some degree of automation in the decision process of physical security deployment, a suitable method of inquiry is needed that appropriately addresses the concerns of those that would use the model or would be influenced by the output of such a model.

BACKGROUND

In designing cost effective perimeter security, Tarr (1994) advocates an integration of both subjective and objective approaches for considering limitations to physical security modelling. His investigation into the modelling of prison facilities, addresses a lack of integration that he describes as a *dichotomy*, suggesting considerable limitations to using either subjective or objective methods in isolation.

The documented approach by Tarr and Peaty (1995) uses a physical security model to build a virtual facility and recognises performance measures such as; *probability of detection, probability of intervention, worst intervention, false alarms, capital costs, and equivalent annual cost.* It can be seen that these attributes of physical security are sometimes difficult to quantify and may lead to the case where calculations are made with large uncertainty. As modelling typically uses quantitative properties for decision analysis it usually has the object system in mind, while greater difficulties are encountered in modelling if consideration is given to the subject system (White, 1975, pp4). Using models for decision making in the presence of large uncertainty can be compounded by the existence of socially constructed variables.

In the initial consideration of a physical security system, Lindfield and Rodger (1992) advise the importance of defining the respective responsibilities of all parties concerned in the procurement system to ensure that each has a clear idea of what is expected of them. They stress formalising of each series of document can provide a history of the installation and protect the commercial and contractual interests of all parties. Some of the documentation specified includes; user requirements, performance specifications, commissioning procedures, system evaluation and routine audits. C*learly, users cannot perform real attacks against their installations in order to ensure that a system is working to specification, and a non-destructive test strategy is needed. To achieve a much higher and more predictable level of performance, users need to be in control of every phase of system procurement* (Lindfield & Rodger, 1992).

To translate the needs of the user into a set of criteria for the system designer, the *Operational Requirement* and the *Performance Specification*, should respectively (Lindfield & Rodger, 1992):
be written in user language
state problems, not solutions
detail all the constraints on the design
And;
Specify performance in terms of measurable attributes
Stipulate how those attributes will be measured, and
Be legally valid and enforceable

Houchin (1999) considers the difficulties in devising a robust costing model from the perspective of a prison governor. He mentions the two spheres of interest, where complexity in pursuing the options for refurbishment (procurement process) is compared to the complexity in options for the individual components of technology. Houchin (1999) further describes the ultimate conclusion of such a model where *the goal of the development is to deliver the full set of operating requirements at the least cost.* He considers the objectives as *likely to be many and complex* and also the cost themselves having a range of components such as: *initial and full-life costs, capital and running costs, fixed and variable costs* where a trade-off in technology and sophisticated equipment will offset staffing costs to a certain degree.

The attributes of an operational requirement provide a basic outline for the input required in any physical security model. The reference to user language and specifying performance in measurable attributes both represent points of interest for modelling in physical security. The suggestion of user language may provide a structure to be used in reducing confusion and describing essential data.

THEORY

This section on theory has been divided into three parts that look to other disciplines to establish a theoretical journey towards a new method of inquiry for physical security modelling and these parts include; Decision theories, Sociology, and Structural Equation Modelling (SEM). There has been an attempt made through each part to demonstrate a common theme of a user language that may be viewed as a defining tool for characterising properties to link across the boundaries of these disciplines.

Decision theories

Decision Theory has had substantial review across many disciplines. As editor for a publication on *Decision Theories in Practice*, White (1975, pp4) is interested in an integrated modelling approach to connect the subject system and object systems in making decisions, where *before one can undertake any decision analysis, one has to consider the questions of 'who are the decisionmakers?' and 'which is the system about which the decisions are being made?' The former group is referred to as the subject system, and the latter is referred to as the object system.*

White (1975, pp4) continues to summarise that where a model can be created as the result of combining several hypotheses, a theory is often a creation of the consequential behaviour of statistics, which can then create a special theory, which may in turn lead to improved decisions or decision processes. On the possibility of modifying a theory, White (1975, pp9) states: *We may seek properties, in a suitable language, which will characterise the actual behaviour, and we may then examine these properties as to their reasonableness.* The reference to suitable language here as a characterising tool for examining the property of reasonableness, may suggest that properties (elements) could be assigned a discrete value state, and from this value state there exists a possibility for language or definitions to be quantified.

For many applications in life, probability measures have been extracted from individuals where response methods have been classified direct if the person responds explicitly (eg placing a number on his subjective belief), and indirect if it is gathered from another mode of response. An alternative to this is assessing a subjective set of probabilities over a set of propositions or assessing a subjective distribution over a range of conceivable propositions (Bunn & Thomas, 1975, pp118).

White (1975, pp190) mentions the Delphi technique as *just a short term analogue of the long term collection, examination and criticism of any knowledge, which takes place over centuries.* According to Ziglio (1996, p3) the Delphi Method is based on collecting and distilling knowledge in a structured manner from a group of experts with controlled opinion feedback. He further describes it as being

used to support decision making and cites Delbecq et al., where group members pool their judgements, inventing or discovering a course of action (Ziglio, 1996, p3).

As a useful tool, a profile analysis could result with a comparison to a subjective distribution over a range of conceivable propositions. This may be developed from a method such as the Delphi technique and may be represented graphically as per the following description by Green et al. (1988, pp107):

If we have n persons measured on m variables, the profile of each person can be portrayed as a point in m dimensions. If we also know the group to which each person belongs, and the groups differ in terms of average profiles, often called centroids, we might expect to find different groups occupying different regions of the space. The less overlap noted among intergroup profiles in that space, the more likely it is that discriminant analysis can help us separate the groups (Green, Tull & Albaum, 1988, pp107).

The importance of this section suggests that a database of n dimensions, comprising language components all suitably defined, could create a reference body of knowledge to be used for decision analysis. Where attitudinal representations such as that by Smith and Robinson (1999, pp32) have demonstrated a graphical relationship (2-D) between pair-wise barrier types for both a novice group and an expert group, similar n dimensional representations could be used as a reference point between a user and an expert, to rank physical security data by a relative value state. This may represent the basis for automating policy decisions for a physical security model.

Sociology

Investigative methods in the behavioural sciences often use the natural sciences for guidance in the selection of method of inquiry. The previously mentioned basis of (user) language provides an alternate method of inquiry into a problem. This section on sociology, also called behavioural science, will introduce Phenomenological Sociology as a method outside the standard objective-subjective debate and attempt to provide a basis for modelling from the perspective of language definitions.

The terms subjective and objective, in sociological debate according to Smart (1976, pp83), have preconceived bias for; *objective connoting true, scientific, rigorous and real, apparent to Any Man, whereas the term subjective by contrast denotes arbitrary, biased, unscientific and personal opinion. Subjective becomes in this sense a term of virtual condemnation, taboo for an aspiring science.* Where the objectivist, subjectivist, and Bayesian would all follow similar steps in; formulating a problem, developing, predicting and testing of the hypothesis, and then test and analyse results, the phenomenologist is opposed to the use of explanatory hypotheses that represent preconceived ideas of the phenomenon and can be viewed as problematic in selective perception and measurement distortion (Green, Tull & Albaum, 1988, pp41).

Smart (1976, pp75) describes externalisation and objectification as taking place primarily through language. His opinion is that Phenomenological Sociology simply deals with accounts, descriptions, conversations and talk. Further to this, the world takes on a factual appearance only after externalising and objectifying occurs, rather than the social world being represented by any defining characteristics of human interaction. Smart (1976, pp75) captures the beliefs of the phenomenologist with a description of the social world as *a linguistic and cognitive world,* where *the task of sociology becomes one of describing the processes by which the social world is constructed through accounts, readings, understandings and interpretive procedures.*

Green et al. (1988, pp41) state the following:

Four steps are recognised by enough phenomenologists to qualify as representative of the approach generally followed:

1. Suspension of prior conceptions
2. Description of the phenomenon

3. Determination of universal elements
4. Apprehending of relationships

In a real sense, this method of inquiry is analogous to what many researchers call a "fishing expedition" without knowing anything about the body of water (Green, Tull & Albaum, 1988, pp41). The metaphor used in this case really seeks to emphasises the importance of the universal elements, as they will identify the important areas of the inquiry and direct any further requirements. The metaphor captures the spirit of Phenomenological Sociology where the lack of any set hypothesis requires that all elements are given the same level of importance prior to the inquiry.

An additional application of Phenomenological Sociology to modelling may be found within the macro micro focus of a society. Smart (1976, pp87) mentions that within sociology *the macro-micro distinction works to allow the sociologist to shift the focus of his investigation or interest from the "hard data" of social structure (macro) to the "soft data" of individual and group experience of social reality* (micro). Smart (1976, pp75) describes the ambition by behavioural scientists of creating a *syntheses of these different "levels" of analysis.* He continues to describe the focus at the macro-level as the *'more objective', scientific and relevant approach to the study of society* and the micro-analysis regarded as secondary, supplementing and small-scale, delimited areas of inquiry (Smart, 1976, pp87).

In the creation of a phenomenological approach to modelling, four recognised points; *suspension of prior conceptions, description of the phenomenon, determination of universal elements and apprehension of relationships* have been addressed in this paper's summary analysis. In line with previous sections, the common link of (user) language is drawn upon again. Smart's (1976, pp87) reference to a synthesis of different levels, echoes the sentiments of Tarr's (1994) dichotomy, and is also in line with those of Hill (1982), as described in the following section.

Modelling

This section on modelling comprises some interpretations from a hierarchical approach, introduces the requirements for structural equation modelling (SEM) and introduces Lisrel.

Hill's (1982) interpretation of modelling combines the process of simplifying and isolating aspects of reality leading to hierarchical descriptions. He establishes each hierarchical level as having its own principles with insight available into other levels depending upon whether one moves across a level or up or down. Also he states that hierarchical descriptions will apply equally well to both structures and functions (Hill, 1982). An example of a hierarchical structure has been included in Figure 1.

Ward (1989) further supports the view that there exists a common tendency to under-emphasise qualitative aspects of a decision situation. In the following quantitative rationale of the requirements he states:

- numbers form a common language for more effective and objective communication;
- decision situations can be precisely defined and ambiguity minimized;
- the logic of the decision process becomes more explicit;
- decisions, and subsequent changes in decisions, can be defended in a rational and consistent manner.

Moskowitz, (1975, pp328) describes the need to begin with a theory as inherent in all modelling and states how a theory can be expressed symbolically in terms of a *formal axiomatic structure or explicitly stated as a set of logical premises.* The components of a model can be found in *protocols from decision-makers, the researchers' experience, previous related research, etc.* (Moskowitz, 1975, pp328).

In general modelling analysis, Green et al. (1988, pp41) provide discussion where modern marketing research has become increasingly interested in the useful application of path analysis and SEM to

establish causal relationships. Reisinger & Turner (2000, pp2) in turn report SEM as seeking to explain patterns of dependence relationships simultaneously between a set of latent (unobserved) constructs, each measured by one or more manifest (observed) variables such as; distance, cost, size, weight or height. The manifest variables can be measured through respondents, via data collection methods or be gathered as secondary data from published sources. SEM expresses linear causal relationships (similar to multiple linear regression) between variables and constructs. Many important marketing, psychological or cultural concepts are latent constructs, with unknown reliability, SEMs can model important latent constructs while taking into account the unreliability of the indicators (Reisinger & Turner, 2000, pp 3).

The discussion of a Lisrel (a commercial SEM) in this paper is to demonstrate the steps that can be used to adapt the SEM analysis to a physical security environment via its constructs and variables. The eight stages listed below are associated with SEM. It is the initial two stages that are of special interest in relation to developing a sound method of inquiry. Both Stage 1 and Stage 2 will be used as guiding concepts in the following section. There will be a cross-comparison made with the four steps of Phenomenological Sociology to ensure that the suggested method of inquiry can be demonstrated as valid for any model simulation. A summary of Lisrel is as follows:

Lisrel stands for LInear Structural RELationships and is a computer program for covariance structure analysis. It is a multivariate technique, which combines (confirmatory) factor analysis modelling from psychometric theory and structural equations modelling associated with econometrics. It was originally introduced by Joreskog and Van Thillo in 1972 ... Lisrel is considered by most researchers as the flagship structural equation modelling (SEM) program (Reisinger & Turner, 2000, pp 2). Lisrel modelling and testing stages have been given as follows:

Stage 1 Development of a theoretical model (with defined causal relationships)
Stage 2 Construction of a path diagram.
Stage 3 Formal mathematical specification of the model
Stage 4 Variance/covariance or correlation (assessment of the sample size)
Stage 5 Model identification (enable parameter estimation)
Stage 6 Assessment of the model fit
Stage 7 Modifications to the model
Stage 8 Cross-validation of the model
(Reisinger, Y. and Turner, L. 2000, pp 6-18)

The adaptation of SEM to physical security may provide a strong basis to calculate values for some of the complex constructs required for decisions making. The fitting of arbitrary mathematical functions to any data set should be based on theory as described above and any measurements to support a specific theory will always need to be independently verified. The definitions of Stage 1 and Stage 2 of Lisrel modelling will be used as a guide for the method of inquiry as proposed.

ANALYSIS

For this section the points of interest established from Phenomenological Sociology and the first two Lisrel requirements are combined for a method of inquiry and are presented for analysis as follows:

1. Suspension of prior conceptions

This initial concept requires that all prior theoretical models be disregarded and the suspension of any path diagrams that currently exist in physical security modelling.

2. Description of the phenomenon

This area requires a theoretical model that describes physical security modelling. The development of a path diagram of real phenomena of physical security will assist in the process to achieve a set of

elements and element relationships and thereafter relative mathematical representations of these relationships.

3. Determination of universal elements

The development of a theoretical model to determine the absolute range of variables for modelling physical security is required. The concept of (user) language can provide a good base for the scope of limitations and the extent of the universal elements (where in a model simple text would represent each element and each element relationship). The theory also requires a testing mechanism to provide feedback on the set of elements proposed. Exploratory studies and questionnaires of a Delphi-type, may provide for the identification and formulation of problems, that can then lead to the identification of all relevant variables. Results from questionnaires can be iterated a number of times for developing more specific determination techniques (Green, Tull & Albaum, 1988, pp497).

Constructing a path diagram for universal element determination is essential so that with any physical path through an environment, each element encountered can be individually described by a model in the same order of interaction as the physical environment. As a possible means of assessment, the physical security model may seek to analyse paths through a universal element database where the same path can be defined through a physical environment. The significance of the path diagram then, is in its parallel to the physical world which will enable users to navigate a physical security model in much the same way that they would navigate a physical space.

4. Apprehending of relationships

The development of a theory to apprehend the relationships between variables is also a requirement to the definition of causal relationships. Green et al. (1988, pp105) discuss the real-world environment that determines the "rules" that describe relationships between events. White (1975, pp13) attempts to provide value determinations of the relative importance of measurable attributes and alternatively suggests that decision makers can assign numbers of importance to elements. The relationship of the user to a physical security model should be evaluated in such a theoretical model using the Delphi Technique as is proposed by White (1975, pp190).

The construction of a path diagram for apprehending of the relationships between elements will serve to be the rationale for paying particular attention to the secondary problems (as per a micro-macro focus) and how they may have an influence on the resolution of the primary problem or provide insights into the best solution (Barron, 1975, pp203). In drawing inferences from causal relationships, three types of evidence are available: *(1) associative variation, (2) sequence of events, and (3) absence of other possible causal factors* (Green, Tull & Albaum, 1988, pp107). This evidence can possibly recognise a path diagram for apprehending physical security data relationships to relate back to the principles of Defence in Depth. The path diagram for the element relationships may show each single element as a possible input function into a profile combination of the user and other elements. Although each profile may contain the same elements, they may yet have different relationships for a model depending on the extent of causality.

EXAMPLE

A practical example of how universal elements can be defined and assigned value states has been included in the following section. The elements can be initially identified through a Phenomenological sociology-type process, and organised through a process hierarchical structure to form a universal element database. Relationships between the elements can be examined through a Delphi-type questionnaire process and finally constructed as a series of path relationships as per requirements of SEM (Lisrel). In this example the elements are initially identified in quotations and the value states identified in Italic font.

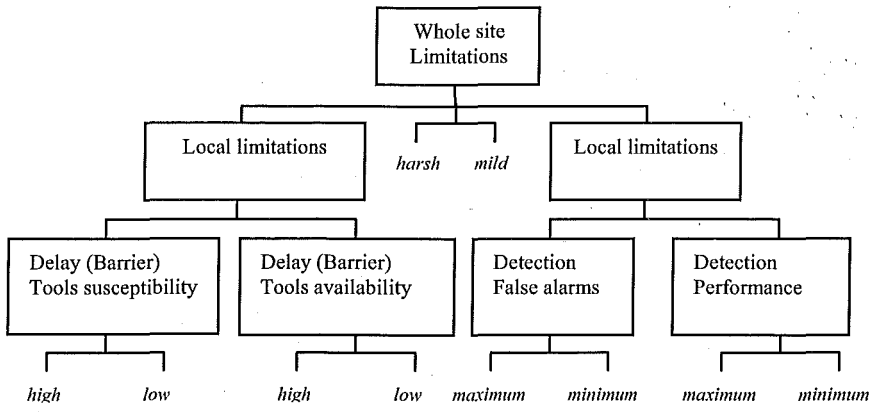


Figure 1. Example of a structural hierarchy for physical security

Limitations for physical security

To estimate the operating limitations for a facility, it is necessary to separate the limitations into "local" and "whole site" categories. The severity of "weather" can be an example of an influencing force. The influence of weather is one that affects the whole site and is therefore independent of any "region" or "time of day". The value state for the states of weather to be assigned to a facility may include *mild* or *harsh*. These can act in conjunction with local limitations to affect both the detection rates and the delay time of barriers. The values may show that in general the more harsh the weather the lower the detection ability for intrusion, and, conversely the more harsh the weather, the greater the delay time of the barriers. These states have been allocated to give a broad yet simplistic overall adjustment to reflect true variance where weather is a influence.

The greatest source of nuisance in operating a Perimeter Intrusion Detection System (PIDS) is the rate of "false alarms". The local limitations may be defined as those that affect the detection "performance" of PIDS for a particular region. A high rate of false alarms leads to two problems; a waste of money and resources in attending false alarms, and the chance that a real detection may be treated as a false alarm. The value state for the false alarm influences of a facility may include *minimum* and *maximum*. The choices can be allocated to allow for adjustment in performance where either; lighting restrictions, animal interaction or other local influences may cause false alarm conditions and reduce the effectiveness of detection.

The greatest source of success in negating a barrier is the "availability" or "susceptibility" of "tools" as a countermeasure. Tools can be legally obtained, stolen or manufactured. The local limitations may be defined here as those that decreasing the delay time of one or more barriers. The value state for the availability (susceptibility) of tools can include *low* and *high*. These states can be allocated to provide adjustment capability in barrier performance and reflect higher risk conditions where appropriate tools can be used against defences. Where tools that are capable of delay reduction may be used close to the immediate region at risk, the value state of the accessibility or susceptibility acting in conjunction with site limitations, can determine the relative effectiveness in delay time of any barrier.

In the practical example, the universal elements have been recognised and value states defined. The value states will be determined by controlled feedback with respect to the policy of the decision makers, the technical knowledge of the security expert and the general opinion of the users of any such system. After initial element-value relationships have been determined for all cases (universal element database), the specific case that relate to a facility can be evaluated as a path analysis to provide a relative estimate of physical security effectiveness against other possible paths and assist the decision process in deployment of physical security.

CONCLUSION

The process of integrating theory from other disciplines may offer an alternative method of inquiry into the process of designing a physical security model. The three parts of theory investigated in this paper included; Decision theories, Phenomenological Sociology, and SEM. The common theme that was suggested to link each of the three parts was that of a suitable (user) language. The significance of this suggests that the method of inquiry may be directed towards a more language based concept.

In the analysis suggesting a method of inquiry, an inter-comparison between the steps of Phenomenological Sociology, and the requirements of Lisrel has been established using Delphi-type questionnaires to identify universal elements and their relationships in physical security modelling. Some of the main points in the analysis were; the necessity of path construction to mirror real world paths through a facility, the determination of universal elements, the evaluation of their relative importance and any causal relationships associated with those elements.

A suitable physical security model could be used by both decision makers and security consultants to evaluate existing structures or proposed changes to existing structures with respect to physical security. The benefit of such a model may be seen as a consistent and fully documented decision structure, with a user interface that allows for the input of appropriate data.

This paper has examined an extension of theory to address difficulties involved with physical security modelling and has suggested an alternative method of inquiry to current treatments. This approach has a potential to provide physical security modelling with a universal element database of both object related and subject related knowledge that may provide assistance in policy decisions for the user.

REFERENCES

Barron, F.H. (1975). An Information Processing Methodology for Inquiring into Decision Processes. In White, D.J. (Ed). *Theories of Decision in Practice*. Hodder and Stoughton, London. pp. 195-206.

Bunn, D.W., and Thomas, H. (1975). Assessing Subjective Probability in Decision Analysis. In White, D.J. (Ed). *Theories of Decision in Practice*. Hodder and Stoughton, London. pp. 117-127.

Green, P.E., Tull, D.S., and Albaum, G. (1988). *Research for Marketing Decisions*. Fifth Edition. Prentice-Hall, International, New Jersey.

Hill, P.W. (1982). *Process hierarchy theory: A holistic approach to theory building in Education*. PhD Thesis, Murdoch University, Western Australia.

Houchin, R. (1999). Reflections on the incorporation of modern technology in the refurbishment of a Victorian prison. *Proceedings of the 1999 International Carnahan Conference on Security Technology*, Institute of Electrical and Electronics Engineers. Pp 412-422.

Lindfield, A.G., and Rodger, R.M. (1992). Fence Detection Systems-Achieving the Desired Performance. Paper for ADPA 1992 Conference. Police Scientific Development Branch, Home Office, U.K.

Moskowitz, H. (1975). Model Choice for Decision Making: Issues and Evidence Regarding the Use of Regression Models of Behaviour. In White, D.J. (Ed). *Theories of Decision in Practice*. Hodder and Stoughton, London. pp. 328-340.

Reisinger, Y., and Turner, L. (2000). Structural Equation Modelling with Lisrel: Application in Tourism. *Working Paper 35/00, June 2000*. Monash University, Faculty of Business & Economics. pp. 1-29.

Smart, B. (1976). *Sociology, phenomenology and Marxian analysis.* Routledge & Kegan Paul, London.

Smith, C.L. and Robinson, M. (1999). The understanding of security technology and its applications. *Proceedings of IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology*, pp. 26-37. Keynote Address. Madrid, Spain.

Tarr, C.J. (1994). Cost Effective Perimeter Security. *Proceedings of the 1994 International Carnahan Conference on Security Technology*, Institute of Electrical and Electronics Engineers. pp. 60-65.

Tarr, C.J. and Peaty, S. (1995). Using CLASP to Assess Perimeter Security. *Proceedings of the 1995 International Carnahan Conference on Security Technology*, Institute of Electrical and Electronics Engineers. pp. 93.

Ward, S.C. 1989. Arguments for Constructively Simple Models. *Journal of the Operational Research Society.* vol 40 pp 141-153.

White, D.J. (1975). The Nature of Decision Theory. In White, D.J. (Ed). *Theories of Decision in Practice.* Hodder and Stoughton, London. pp. 3-17.

Ziglio, E. (1996). The Delphi Method and its Contribution to Decision-Making. In Adler, M. and Ziglio, E. (Eds). *Gazing into the Oracle.* Jessica Kingsley Publishers. London. pp3-33.

Exploiting sshd1 with logarithmic complexity

K. Lyytikäinen[1], P. Korpinen[2], T. Virtanen[3]

[1] *Helsinki University of Technology, Finland,*
E-mail: kalle.lyytikainen@hut.fi;

[2] *Helsinki University of Technology, Finland,*
E-mail: pekka.korpinen@hut.fi;

[3] *Telecommunications Software and Multimedia Laboratory*
Helsinki University of Technology, Finland,
E-mail: tpv@tml.hut.fi;

Abstract

Ssh is a widespread method for securing sensitive connections and data transfer. Ssh1 servers introduced a detection mechanism against CRC32 compensation attack, which includes a design fault. It enables remote hosts to write arbitrary values into the server's memory. With carefully engineered cipherpackets, an attacker can investigate the server's memory using search methods such as binary search. With such methods, the attacker can find exact addresses for critical stack frame and program variable memory addresses. Knowledge of these addresses makes executing malicious program code trivial. Binary search reduces the required time to breach the victim system to reasonable levels.

Keywords: *ssh, sshd1, exploit, buffer overflow,*

INTRODUCTION

Ssh is widely used in remote administration and distributed systems. There is a great number of servers in the Internet and corporate networks that are administered by systems specialists. These specialists are often concentrated in certain parts of organizations and do most of their system administration remotely. Ssh is perhaps the most used method for securing remote logins and sessions to computer network servers.

Many business systems are distributed on separate servers and thus need to communicate with each other. PKI and ssh can be used to encapsulate that sensitive communication into a meaningless ciphertext. Web servers, database servers, database mirrors, e-mail servers and such often use ssh tunnels to communicate with each other securely.

Majority of business-scale systems come equipped with ssh as default. If an intruder can operate a foreign server remotely relatively easily, it is a great threat to majority of servers on the net. Exploiting this vulnerability usually grants the intruder super user access to the system and thus access to all data and configurations. Sensitive information can be compromised and servers can be sabotaged and used for further attacks, such as DDoS. Using this kind of exploit application can be easy to use – the attacker does not have to supply much more than target host address. It also does not require much network bandwidth to operate and thus can be used more widely Solar Designer, Dug Song (2001), Securityfocus.com (2001c).

Exploiting the vulnerability with classic brute force methods would require millions of malicious IP-packets to be sent. This would take too long and the intruder would be spotted before successful breach with very high probability. Reducing the exploit complexity to logarithmic time, even partially, makes the exploit quite swift and increases the chance to gain root access significantly. It is possible to breach the victim system with only some tens of malicious packets.

EXPLOIT BASICS

Introduction

The crc32 compensation attack detector in various ssh versions has a bug that allows a malicious attacker to write to memory locations on the server. If the packets are constructed with sophistication, they may allow foreign code to be executed on the server side under the root account. Vulnerable sshd versions are listed in Securityfocus.com (2001a). Blackshell (2002) has developed a script that can detect vulnerable servers.

When the client or the server receives an encrypted packet, it checks if the packet is tampered using the crc32 compensation attack detector. Unfortunately most of ssh1 daemon implementations suffer from the vulnerability in the detector. The ssh protocol is presented in Ylönen (1995).

Memory layout

The operating system's memory layout is in a crucial role in this exploit. Knowledge of the target architecture is required. Figure 1 represents the general memory layout of a user process in linux Johnson (1996):

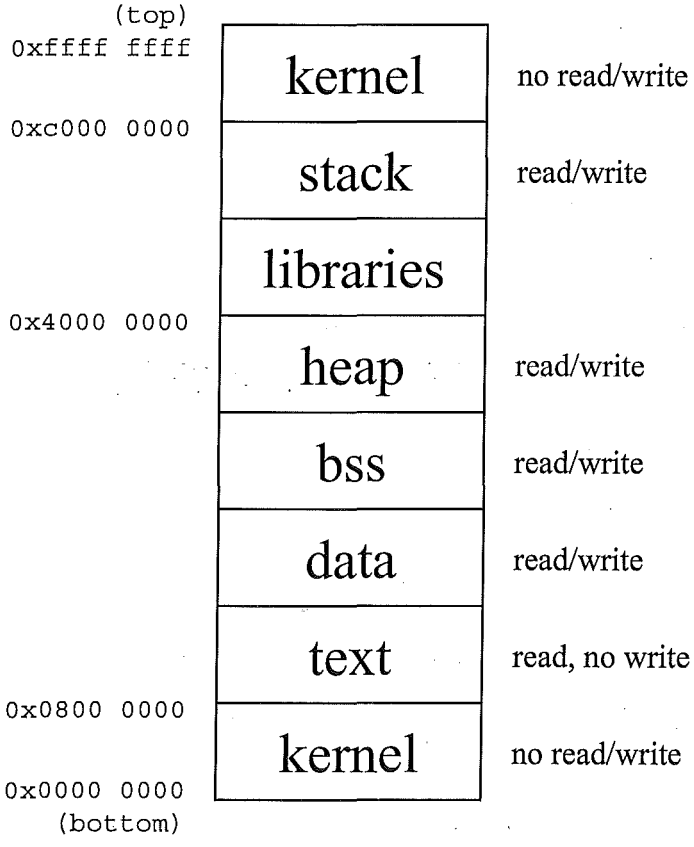


Figure 4: Linux memory layout

Vulnerable code and its execution

The exploit is based entirely on the bug in detect_attack() function in source file deattack.c. The function detects crc32 compensation attacks, but it also introduces another security threat.

Critical parts of the detect_attack() function:

```
int detect_attack(unsigned char *buf, u_int32_t len, unsigned char *IV)
 {
   ..
        static u_int16_t *h = (u_int16_t *) NULL;
        static u_int16_t n = HASH_MINSIZE / HASH_ENTRYSIZE;
        register u_int32_t i, j;
        u_int32_t l;
   ..
#1  for (l = n; l < HASH_FACTOR(len / SSH_BLOCKSIZE); l = l << 2)
                ;

        if (h == NULL) {
                debug("Installing crc compensation attack detector.");
#2              n = l;
                h = (u_int16_t *) xmalloc(n * HASH_ENTRYSIZE);
        } else {
   ..
        for (c = buf, j = 0; c < (buf + len); c += SSH_BLOCKSIZE, j++) {
#3              for (i = HASH(c) & (n - 1); h[i] != HASH_UNUSED;
                     i = (i + 1) & (n - 1)) {
                        if (h[i] == HASH_IV) {
                                if (!CMP(c, IV)) {
                                        if (check_crc(c, buf, len, IV))
                                                return (DEATTACK_DETECTED);
                                        else
                                                break;
                                }
#4                      } else if (!CMP(c, buf + h[i] * SSH_BLOCKSIZE)) {
                                if (check_crc(c, buf, len, IV))
                                        return (DEATTACK_DETECTED);
                                else
                                        break;
                        }
                }
#5              h[i] = j;
        }
        return (DEATTACK_OK);
}
```

The function is called after each received ssh packet. The buf argument is the buffer where incoming (encrypted) packet's payload is located, 'len' is the length of the buffer The 'IV' is always NULL.

The problem resides on the #2 where a 32-bit integer ('l') is assigned on a 16-bit integer ('n'). The 'l' is calculated on #1, and is dependant on the 'len' variable. If the 'len' is bigger than about 88000, the 'n' is set to zero.

Then 'h' is allocated with call to malloc and the requested becomes to zero because 'n' equals 0. The allocated size is the minimum allocation block.

The outer for-loop (near #3) goes through the whole packets. The variable 'j' is the block number (blocks are 8 bytes long). On #3 the HASH()-macro gets the first 32-bits from the current block. Because the 'n' is zero the (n-1)-clause is evaluated as 0xffff. So, the variable 'i' has the first 32-bits of the current block.

On #5 is the assignment operation. This is where the memory can be overwritten. Variable h is the pointer to 16-bit values. So the h[i]=j clause can be expressed also with C-style byte pointers: *(h+2*i)=j. Thus, the value of 'j' (the current block number) is written to a specific distance from h. Effectively, the index of the block in buf denotes the value written to memory and the value of the block in buf denotes the offset from h to be overwritten.

Challenges

As one can see after examining the detect_attack() function, the hardest problem in the exploit is finding the addresses of critical local variables and the function return address.

A successful exploit must investigate the following:

- distance between h and buf
- distance between h and the return address on the stack
- the value of the return address
- rough absolute address of buf

With these parameters the exploit is somewhat straightforward. Educated guesses can be made based on knowledge of the host system. This information can be easy to gather using appropriate tools. The attacker could investigate software versions remotely by querying the version strings of various services on the host and comparing them to operating system distribution defaults.

Memory Layout Investigation Process

Exploitation parameters can be determined by sending numerous packets to the server. An address can be found by sending a packet that writes to a specific region of memory.

Deattack.c can be induced to write to certain memory locations, should writing to them be allowed. If the server dies (client sees the connection closing), the algorithm should change the address and start again. The server dies because of segmentation fault (i.e. writing to a read-only memory). If the server reports "invalid bytes on packet" the writing was successful and a different region was hit.

The authors created an exploit application that mimics some aspects of the shack implementation. It uses the modified ssh client to try connections to the victim, with certain intelligence. The application generates malicious packets for the victim and lets the modified ssh client send them. Depending on the output of the server, the application makes conclusions on the effects of the malicious packet. The server may return such messages as 'corrupt bytes in packet' or 'CRC32 compensation attack detected'. Although, the most frequent response is a mere connection close since the server crashes often (SEGV). These output strings are analysed and next malicious packets are engineered with that analysis. It is very typical that the connection closes after the first cipherpacket, since it is somehow malformed. For example our exploit does not care about CRC checksums and thus the sshd closes the connection. Although, a lot can happen before the connection is closed...

Reverse engineering

The authors reverse-engineered an implementation of this exploit. The implementation is called 'shack' Team TESO (2001). Team TESO is believed to be its creator. The shack exploit surprised us with its capability to find out critical addresses concerning the exploit Dittrich (2001), Incidents.org (2001), Lee (2001). A utility program called 'tcpflow' and GNU debugger (gdb) were used to reverse engineer the operation of the shack exploit Elson (2001), Visscher (2001).

Gdb was used to diagnose various erroneous states and in confirming memory address calculations. However, the usage of the debugger was hindered. While running sshd under the debugger, addresses of variables were different from running without the debugger.

Finding distance to 'buf'

The first thing that can be calculated in the server's memory space is the distance from variable h to variable buf (according to detect_attack() function). Shack implementation uses 'the TESO method' to investigate this distance. In the authors' implementation, this is the phase that is done in logarithmic complexity Team TESO (2001).

The following format of cipherpacket was used to search for the distance from h to buf:

Extract from packet #1:

```
00000b8 00 00 00 00 ff ff ff ff 00 00 00 01 ff ff ff ff
00000c8 00 00 00 04 ff ff ff ff 00 00 00 05 ff ff ff ff
00000d8 00 00 00 08 ff ff ff ff 00 00 00 09 ff ff ff ff
00000e8 00 00 00 0c ff ff ff ff 00 00 00 0d ff ff ff ff
00000f8 00 00 00 10 ff ff ff ff 00 00 00 11 ff ff ff ff
...
```

This is the first type of packet that the exploit uses. It consists of 184-byte header and 102400 bytes 8-byte blocks that consist of a small 32-bit number and 0xFFFF. Consecutive packets increase the small numbers. When sshd processes this kind of packet, it will set i to first 32-bits of each 8-byte block (#3 in detect_attack()). It tries to read memory at h[i] on the same line. This read will cause SEGV if i is such an offset to h that the address is not readable. SEGV also happens if buf+h[i]*8 is unreadable. With this kind of packet, the sshd will practically always SEGV at #4 in detect_attack() when done thousands of times in a row.

The trick is to have i[small number] point always to 0xFFFF. This is same as 'HASH_UNUSED' in deattack.c, thus the memory read is skipped. The packet contains 12800 small numbers that make h[i] point to a loose array of approximately half the length of the packet. When the first small number in the packet points to the first half of the buffer, each h[i] will have value 0xFF and the memory read is skipped. It would be extremely rare to find such an array in memory in a place other than buf. The memory is first scanned with half-packet length increments until the set of small numbers survive the execution. Next, binary search is used to find the smallest small number that will not cause SEGV. A hit to buf will cause 'corrupt bytes' response from the sshd. This binary search will reduce the order of magnitude of required packets to find this variable successfully. Average number of brute force packets is 6400 (12800/2) for this variable (h-buf). The binary search reduces the number of packets approximately to 14 (log2 12800), or 1/470 of 6400.

In the first phase, the small number is increased by packet_length / 4 so that its pointed address h[small number] points packet_length / 2 forward on each packet. This is a fast scan through the memory to find approximate distance from h to buf. The second phase is a binary search, which will find the distance accurately and reliably.

Finding distance to stack frame

The next memory address to find is h-to-stack frame distance. The stack frame is of interest since the saved EIP of the parent function is there. The authors have no exact knowledge of the methodology how shack implementation does this. The authors' implementation tries to find the h-to-kernel space distance since the stack frame is just below the kernel boundary (0xC0000000).

In this phase, a dedicated type of ssh-cipherpacket is used to test if h[i] is readable (#3 in detect_attack()) and small enough that buf+h[i]*SSH_BLOCKSIZE is readable (#4 in detect_attack()). The stack area contains such small numbers. Only one address is tested and the process is terminated as quickly as possible. The pattern of '0x00002860 0100FFFF's is a telltale sign for the sshd that this packet uses CRC32 compensation attack. We want the sshd to think so in order to close the connection right away and speed up the investigation process.

The authors' exploit is a simplified version of the 'shack' implementation in this phase. Authors' current implementation brute forces the distance beginning with an educated guess. It starts testing addresses deep in the kernel space and comes down gradually and finds the lower bound of the kernel space. With this distance, authors' implementation can calculate the addresses for h and buf.

Shack uses three different packet types to hunt down the h-to-stack distance with binary search. It relies on the fact that the buffer content is partially stored in the stack, probably in a local variable.

Executing foreign program code on the server

The heart of the exploit is to send the packet that finalizes the breach to the victim host. This packet combines basically three important functions: It rewrites the saved EIP in the stack frame, exits the attack_detect() function call as quickly as possible, and when returned, executes the shellcode that lets the attacker telnet to the victim and control a root shell Aleph One (2002), jsb4ch (2002).

The first part of the packet is designed so that the for loop (before #3 in detect_attack()) increases the value of j. These blocks are filled with offsets for h to point to the following 0xFF's. This will make the for loop run quickly and smoothly until j is what we want it to be. We want j to be that value which will be written to saved_EIP. The least significant word can be whatever. Since detect_attack() will try to read memory at buf+h[i]*SSH_BLOCKSIZE at #4, the value written on top of the old word in the memory cannot be very large. This means that only small numbers can be written to memory with detect_attack(). This is a reason for only writing the most significant byte of the saved_EIP. However, this is no problem for the exploits, because the NOP sled can be over 0xFFFF long and it does not matter what the LSW of EIP is if the MSW is correct.

When j is the desired saved_EIP_MSW, i is set to the h-to-saved_EIP offset. Now detect_attack() will have to reach point #5 to write the new EIP with correct values of i and j. This is only possible when if clause at #4 is true and the CRC will be checked with check_crc(). This check will not generate 'CRC32 compensation attack detected' message since there is no attack pattern. On the contrary, the crc_check() will pass and the for loop is broken out of. Now h[i] will be j and detect_attack() has been exploited.

One more thing: The comparison at #4 will only pass if the block at c is the same as buf+h[i]*SSH_BLOCKSIZE. j will be the target saved_EIP_MSW (such as 0x807) and c is buf+j*SSH_BLOCKSIZE, therefore the contents at buf+h[i]*SSH_BLOCKSIZE need to be the same as at c. h[i] hopefully points to the MSW of saved_EIP and has a value that is the original return address for detect_attack() at this point. Once again, these blocks are followed by the CRC32 compensation attack pattern for exiting the outer for loop as quickly as possible.

CONCLUSIONS

The shack-exploit was first used successfully in fall 2001, even though the vulnerability was reported in February that year Rafail, Dougherty (2001), SSH Communications Security (2001), Cisco (2001), Starzetz (2001), Lanza (2001). This reflects how slowly system administrators updated ssh daemons and how complex this vulnerability is to exploit. Security experts assumed this ssh1's vulnerability to be so hard to exploit that it were no serious security threat. History shows that shack was able to break though this weakness with astounding sophistication. Detailed analyses of the shack have never before been released.

The reverse engineering presented in this paper reveals some of the tactics and methods used by the shack implementation. The shack method relies on the return values of the ssh daemon. With few different return values, the daemon reveals its internal errors to the client side. The speed advantage of shack is ground braking when compared to brute forcing. This kind of intelligent exploiting makes

braking into remote systems possible in practise and could probably be used on vulnerabilities in other programs as well.

In early 2001, majority of ssh daemons in business systems were running vulnerable versions in regard to this bug. With this sort of exploit application, malicious attackers could have taken control over a major proportion of business and other system infrastructures.

Defending against these type of attacks is complicated. The information sent to clients in most cases is necessary for them to operate rationally. Normally, one can't reduce the amount the information sent back but can make the information less revealing. The rule of thumb is to inform the client as little as necessary about the actions done at the server side. Login schemes are a good example: the system usually will not tell whether the login name or the password is invalid.

References

Aleph One (2002). Smashing The Stack For Fun And Profit. *Phrack ...a Hacker magazine by the community, for the community....* [Internet] 7(49)
Available from: <http://www.phrack.com/phrack/49/P49-14> [Accessed March 20, 2002]

blackshell@hushmail.com (2002). blackshelltool1 - sshd vulnerability scanner. *Securityfocus bugtraq - vuln-dev archive* [Internet] Available from: <http://online.securityfocus.com/archive/82/247801> [Accessed March 20, 2002]
Cisco (2001). *Multiple SSH Vulnerabilities.* Cisco Security Advisory [Internet] <http://www.cisco.com/warp/public/707/SSH-multiple-pub.html> [Accessed March 20, 2002]

Dittrich, David A. (2001). *Analysis of SSH crc32 compensation attack detector exploit* [Internet]. Available from: <http://staff.washington.edu/dittrich/misc/ssh-analysis.txt> [Accessed March 20, 2002]

Elson, Jeremy (2001). *tcpflow -- A TCP Flow Recorder* [software, Internet] <http://www.circlemud.org/~jelson/software/tcpflow/> [Accessed March 20, 2002]

Heinisuo, Rami et al. (1997). *Elektronisen viittaamisen opas* [Internet] Jyväskylä, University of Jyväskylä 1997 Available from: <http://lib.hut.fi/Elehdet/Elviira/> [Accessed January 27, 2002]

Incidents.org (2001). *SSH Attack Activity Update -- CRC32 Exploit "x2" Recovered.* Handler's Diary, Thursday, December 13th 2001 [Internet] <http://www.incidents.org/diary/diary.php?id=118> [Accessed March 20, 2002]

Johnson, Michael K. (1996). Linux Memory Management Overview. *Linux documentation project* [Internet] Available from: <http://www.linuxdoc.org/LDP/khg/HyperNews/get/memory/linuxmm.html> [Accessed March 20, 2002]

jsb4ch@hotmail.com (2002). *ANTI-prym/h4g1s portshell code* [Internet] <http://www.cotse.com/sw/linux/portshell.txt> [Accessed March 20, 2002]

Lanza, Jeffrey P. (2001). *Vulnerability Note VU#945216* [Internet] Carnegie Mellon Software Engineering Institute. Available from: <http://www.kb.cert.org/vuls/id/945216> [Accessed March 20, 2002]

Lee, Rob (2001). *SSH CRC Exploit analysis* [Internet] <http://www.incidents.org/papers/ssh_exploit.pdf> [Accessed June 28, 2002]

Rafail, Jason A., Dougherty, Chad. (2001). *CERT® Advisory CA-2001-35 Recent Activity Against Secure Shell Daemons* [Internet] Carnegie Mellon Software Engineering Institute. Available from: <http://www.cert.org/advisories/CA-2001-35.html> [Accessed March 20, 2002]

Securityfocus.com (2001a). List of vulnerable server versions. *Securiryfocus.com bugtraq list* [Internet] Available from: <http://online.securityfocus.com/bid/2347> [Accessed March 20, 2002]

Securityfocus.com (2001b). *Possible OpenSSH DoS Attack* [Internet] <http://www.securityfocus.com/cgi-bin/archive.pl?id=82&start=2002-01-26&end=2002-02-01&threads=0&mid=004401c181d1$2b91adc0$0400a8c0@pi> [Accessed March 20, 2002]

Securityfocus.com (2001c). *ssh password brute forcing.* [Internet] <http://www.securityfocus.com/archive/82/252405> [Accessed March 20, 2002]

Solar Designer, Dug Song (2001). *OpenSSH subject to traffic analysis* [Internet] <http://www.openwall.com/advisories/OW-003-ssh-traffic-analysis.txt> [Accessed March 20, 2002]

SSH Communications Security (2001). *SSH statement regarding the vulnerability of SSH1 protocol* [Internet] <http://www.ssh.com/products/ssh/cert/> [Accessed March 20, 2002]

Starzetz, Paul (2001). *sshd1 exploit demonstration* [Internet article] Available from: <http://packetstorm.widexs.nl/0102-exploits/ssh1.crc32.txt> [Accessed March 20, 2002]

Team TESO (2001). Shack exploit. *Packetstorm exploit archive* [software, Internet] <http://packetstorm.widexs.nl/0201-exploits/cm-ssh.tgz> [Accessed March 20, 2002]

Visscher Paul (2001). *readelf man page* [Internet] Available from: <http://www.gnu.org/manual/binutils-2.10.1/html_chapter/binutils_14.html> [Accessed March 20, 2002]

Ylönen, Tatu (1995). *The SSH (Secure Shell) Remote Login Protocol* . SSH: The Secure Shell The Definitive Guide [Internet] <http://www.snailbook.com/docs/protocol-1.5.txt> [Accessed March 20, 2002]

Dynamic management of core ad hoc networks

Catharina Candolin and Hannu H. Kari

Laboratory for Theoretical Computer Science
Helsinki University of Technology, Finland,
Email: Catharina.Candolin, Hannu.Kari@hut.fi

ABSTRACT

To cope with infrastructure warfare attacks, it must be possible to rapidly and securely replace the most crucial network components to enable at least some level of communications. In this paper, we present the concept of core ad hoc networking for establishing infrastructures in a dynamic fashion, and a context aware management model to enable nodes to automatically adapt to new tasks in a new environment. By combining the management model with core ad hoc networking, it is possible to replace partially or completely destroyed communication infrastructures and thus allow at least a limited degree of communications to continue.

INTRODUCTION

Infrastructure warfare is a serious threat to society, as people and organizations rely more and more on the serviceability of communication networks. In case the infrastructure is partially or completely destroyed or paralyzed, it must be possible to rapidly and securely reestablish the infrastructure to provide at least a limited degree of communications. Rebuilding the whole infrastructure typically requires time and resources that are not available directly after an infrastructure warfare strike, but some organizations (e.g. the military) may still need a communication network to transmit important information. By rapidly replacing the destroyed infrastructure components with any other available components, it is possible to establish a temporary infrastructure to provide the most crucial services. The replacing components may not have the same capacity as the destroyed components and cannot thus provide the same level of service, however, the most crucial communications can still be maintained.

The main objective of this paper is to introduce the concept of core ad hoc networking as a means to cope with infrastructure warfare attacks. When replacing the components of the communication infrastructure with other components on the fly, it must be possible to ensure the security of the infrastructure by controlling which components may be assigned the given tasks. The components that are assigned to these tasks must in turn be able to rapidly adapt to their new environment. In this paper, we thus focus on security and management issues that are involved in reestablishing the infrastructure.

BACKGROUND

In Figure 1, a simplified view of a communication infrastructure is depicted. The infrastructure consists of the core network and the access networks that are attached to it. The main purpose of the core network is to interconnect the access networks. The access networks in turn offer communication services to end nodes, such as computers, mobile phones, or other devices.

The core network consists of the interconnecting routers and a wide variety of other components. Typically, the core network is fairly static. Components may be added or removed, but they never

move within the network. Trust relationships are static in the sense that components do not need to adapt to changes in the network and dynamically update their trust relationships accordingly.

Access networks are seen to the core network as subnets, represented by the access gateway. The details of their internal structure are, however, hidden from the core network. Access networks contain at least one gateway router and a variety of other components, such as access points, to which end nodes may connect. Access networks may be more dynamic than core networks; the end nodes may move within the network and even the whole access network may change its point of attachment to the core network.

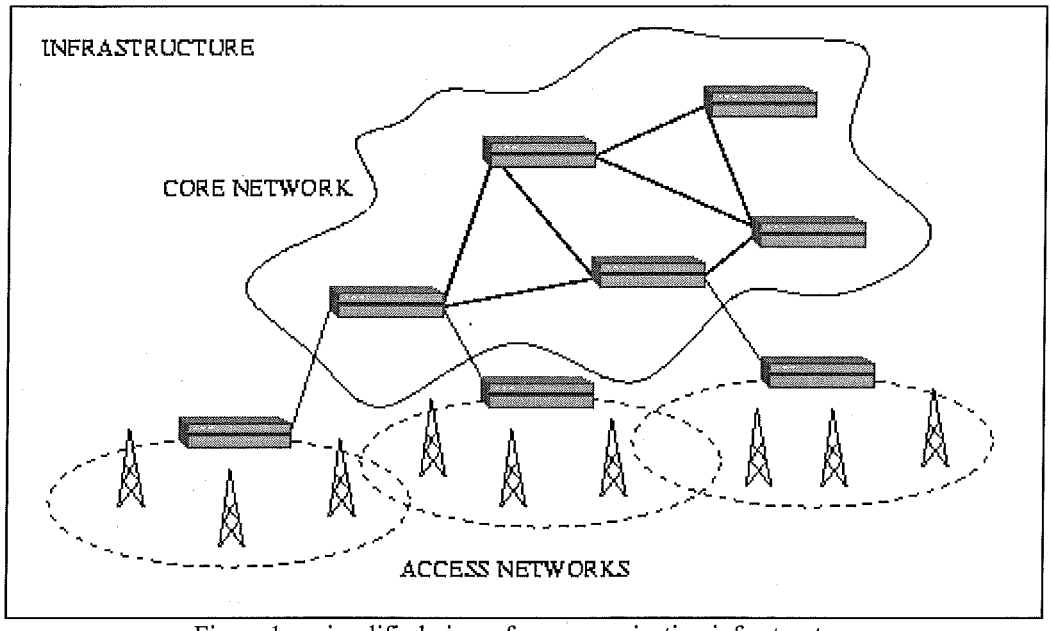


Figure 1: a simplified view of a communication infrastructure

The most dynamic variant of access network is the ad hoc network. An ad hoc network is a collection of nodes that do not need to rely on a predefined infrastructure to establish and maintain communications. Ad hoc networks are thus often considered to be infrastructureless. Most or all nodes participate in network operations, such as routing and network management, depending on their capacity. The main characteristic of such networks is that they are dynamic, and thus able to rapidly adapt to their environment. For example, nodes may enter, leave, or move within the network on frequent basis.

Core ad hoc networks

A *core ad hoc network* is a core network that is established in an ad hoc fashion by assigning the tasks of a core network node to any available node, regardless of where in the infrastructure the node is located. The core ad hoc network thus establishes and maintains the communications without having to rely on any predefined infrastructural components. The core ad hoc network nodes participate in the network operations depending on their capacity, and are able to adapt to changes in the environment. For example, if further attacks are launched against the infrastructure, the remaining nodes dynamically adapt to the situation.

When replacing core network components with arbitrary ad hoc nodes, it must be possible to verify the legitimacy of ad hoc nodes to prevent unauthorized entities from damaging the network any further. Establishing trust relationships between core network nodes and ad hoc nodes is not an easy task, especially since the ad hoc nodes may originate from a different organization than the core network nodes. Furthermore, ad hoc nodes should only be allowed to participate in core network functions if there is a crisis that requires it. Thus, it must be possible for the core network to specify a policy that determines when the scope of the crisis requires external assistance. Ad hoc nodes, in turn,

must have a means of proving authority to provide such external assistance, either by storing a certificate that has been issued before the crisis, or by obtaining a certificate on the fly when the need arises.

The ad hoc node, in turn, must be able to dynamically change its mode of operation to serve as a core network node and to rapidly adapt to its new environment. For example, the node may have to change the protocols it is using. A typical case would be where the node has to change from one routing protocol to another. To perform such changes, a context aware management system is needed.

Context Aware Management

In Candolin et. al (2002), a Context Aware Management (CAM) architecture is introduced. CAM was originally designed to optimize the behaviour of mobile nodes as they change their point of attachment to the infrastructure. When such changes occur frequently, it must be possible for the node to rapidly adapt to the new environment. If the reaction is too slow and the management decisions are made in an suboptimal fashion, the node may end up choosing an uneconomical connection or suffer from bad quality of service.

Traditionally, each application or protocol tries to adapt to the new environment independently. In some solutions, an application is specifically tailored to communicate with one other protocol layer. For example, the Real Time Protocol (RTP) (Schulzrinne et. al 1996) and its control protocol RTCP monitor the quality of the connection and informs the application (e.g. the multimedia video player) about degradation in quality. The application then changes the video encoding to better suit the current quality of the connection. Another example is a combination of Mobile IP (Johnson et. al 2002) with a mechanism for choosing the access medium considered optimal for the applications.
The main problem with such solutions is that the decisions made may be optimal only for the application, but not for the node as a whole. Furthermore, tailoring applications and protocols to intercommunicate adds significant complexity to system design.

To overcome these problems, a new Context Aware Management (CAM) layer is added to the Internet protocol (Deering et. al 1998) stack (see Figure 2). The purpose of CAM is to monitor the environment for changes and to adapt the behavior of the node to the current environment. The applications and protocols need not be aware of the environment at all, but rather focus on taking care of the tasks they have been designed for in the first place. For example, a routing protocol is responsible for establishing routes and forwarding packets, but need not handle issues regarding the choice of access medium. The decisions on how to change the behavior of the node are made by the Policy Manager (PM), which is a centralized entity of the node. The PM is aware of all the modules in the node as well as of their current state. Based on some specific rules (policy) and the information provided by the modules about the environment, the PM is able to make decisions that are optimal for the operation of the whole node, not just single applications. The information provided to the PM by the modules is stored in a common database. This database can be accessed by both the PM and the individual modules; thus, if modules need to intercommunicate, they can do so through the standard interface provided by CAM.

Thus, the CAM architecture contains the following components:

(1) The Context Aware Management (CAM) layer.
(2) The Policy Manager (PM)
(3) The common database (CDB)
(4) The modules attached to CAM.

The purpose of CAM is to provide a common layer to all modules that operate in the node to allow the node to behave in a manner optimal to the current environment. CAM offers two interfaces; one to the modules and one to the PM.

The PM is responsible for making the decisions regarding how the behavior of the node should be changed. The PM is aware of all modules that are loaded into the node. The PM also maintains the state information of each module. Thus, it is possible for the PM to make complicated decisions regarding the functionality of the node. For example, if the node enters an ad hoc network, the PM makes the decision regarding which routing protocol to use and with what parameters. If the node changes to another ad hoc network that uses another routing protocol, the PM may switch off the old protocol and switch on another. Another functionality provided by CAM is event handling. A module may request the PM to send a wake up signal upon the occurrence of a given event. For example, an application may request the PM to signal it once a given QoS level can be offered. This may occur when a network interface (e.g. a WLAN driver) informs the PM of a base station with sufficient signal strength. The security management module of the node may have declared that the given base station is on the list of trusted base stations and access could thus be allowed. The PM then informs the mobility management protocol to make a location update through the given base station. Once the connection is established and the required level of QoS can be offered, the PM informs the application.

The CDB contains all information that is common to several modules. In principle, the CDB is itself regarded as a module. Access to the CDB is managed through CAM and controlled by the PM.

The modules are the protocols, applications, device drivers, and other pieces of software that communicate with each other and the PM via CAM. Modules are organized in a hierarchical fashion so that e.g. all network modules are organized under the category "access devices" etc. Thus, the PM is able to recognize new modules without modification.

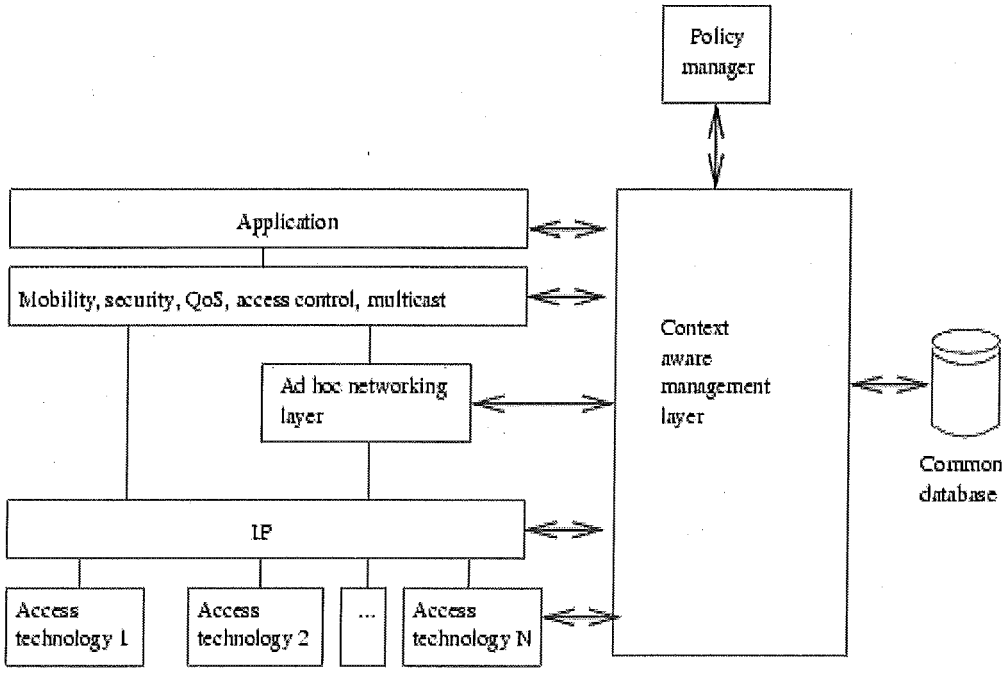


Figure 2: the Context Aware Management architecture

SOLUTION

To cope with infrastructure warfare attacks, we combine the core ad hoc networking idea with the context aware management model. The main idea is thus to replace destroyed core network nodes with ad hoc nodes, which adapt to their new mode of operation automatically by relying on CAM.

Upon destruction of the core network, the ad hoc nodes must somehow realize that they need to change their mode of operation. This can be achieved by self-adaptation, that is, the policy of the node states that there appears to be a crisis and the node should adapt to it according to some instructions, or

by remote reassignment, that is, some other party requests or commands the node to enter the new mode of operation. The details of this process are omitted from this paper.

The main issue when changing the mode of operation is security. First of all, the node must be able to verify that the command to switch to a new mode was issued by a legitimate party (in case of remote reassignment). Second, the remaining core network must be able to verify that the node offering its services is indeed authorized to do so. Performing such verifications can be done, for example, by using digital certificates, such as X.509 (CCITT 1988) and SPKI (Ellison 1999).

A SPKI certificate contains five security related fields: *I,* the issuer of the certificate, *S,* the subject of the certificate, *D*, the delegation bit which denotes whether the certificate may be delegated further, *A,* what the certificate authorizes, and *V,* the validity period of the certificate. In a SPKI based authorization certificate architecture, an issuer may sign an authorization certificate to grant the subject some specific permissions. The subject may, if the delegation bit is set to true, delegate the permissions or a subset of them further by issuing a certificate for the principal it wishes to authorize. This creates a so called certificate chain. The length of the chain is dependent on the values of the delegation bits; as long as the value is set to true, the certificate may be delegated further. Certificate chains represent open arcs beginning from the verifying party, i.e. the issuer, and ending at the party claiming authority. In (Lehti et. al 1998), the concept of forming loops by closing the arcs is introduced. The arcs may be closed by using an online authentication protocol where the claimant proves possession of its private key to the verifying party. If verification succeeds, then access to the resource is granted.

For example, the government may have the authority to take advantage of all available infrastructure in a country in case of a crisis. In such cases, the management privileges of the infrastructure are transferred from the infrastructure owner to the government. The infrastructure owners (*I*) thus issue a certificate stating that the government (*S*) is authorized to fully control the network resources (*A*) for a certain period of time (*V*), and that the government may delegate these privileges further (*D*). The government may now, in turn, delegate the some or all of the privileges further to other organizations, such as the military, which in turn may delegate e.g. a subset of the privileges to its own nodes. When a node is assigned to assist in maintaining the network operability of the core network, it presents the core network with the certificate granted to it. The core network can now verify that the certificate is valid by checking the signatures in the certificate chain, and thus allow the node to access the network.

When the node enters the core network, it must immediately adapt to the new environment. The CAM layer changes the modules (protocols, device drivers, etc.) of the node to fit the new operation mode. For example, the node may switch from an end-node mode to a routing mode and start using an appropriate routing protocol. The rules of the Policy Manager are updated to enforce the management policy of the core network to the operation of the node. For example, if the node has previously adapted a policy where it does not provide routing services to other nodes and now is assigned the task of functioning as a router, the policy of the node will be modified to forward all legitimate packets, and not to engage in any other (previous) actions.

CONCLUSION

Infrastructure warfare is a serious threat to the society, as people and organizations rely more and more on the serviceability of the communication networks. However, if the network is partially or completely destroyed, it must be possible to rapidly and securely reestablish the network to provide at least a limited degree of communications. This may enable at least the most crucial services to continue functioning to some extent.

In this paper, we introduce the concept of core ad hoc networking as a possible solution for rapidly building up a temporary core network until the original network can be reestablished. The main philosophy is to reassign available network nodes to new network functions in an ad hoc fashion.

Security of reassignment is achieved using certificates, so that the core network can verify that the node is authorized to perform the tasks given to it, and the node can verify that the command to switch to the new functionality is legitimate.

When the node enters its new operational mode, it must be able to rapidly adapt to the new environment. This may include, for example, changing the access technology or routing protocols used. To cope with such changes, the node implements a Context Aware Management architecture, which monitors the environment and makes decisions regarding the behavior of the node.

By combining the concept of core ad hoc networking and CAM, it is possible to provide at least a certain degree of connectivity in case the core network is partially or completely destroyed. Although a network established in this manner is not necessarily able to provide the same capacity as the original network, it is still possible to maintain communications until the network can be completely rebuilt.

ACKNOWLEDGMENTS

This research was funded by the Finnish Defence Forces.

REFERENCES

Candolin C. and Kari H.H. (2002) *Context Aware Management Architecture,* IETF Internet Draft <draft-candolin-cam-00.txt> (work in progress), June 2002

Deering, S. and Hinden, R. (1998) Internet Protocol, Version 6 (IPv6) Specification. Request for Comments 2460, IETF, 1998

Ellison, C. (1999) SPKI Requirements, Request for Comments 2692, IETF, 1999

International Telegraph and Telephone Consultative Committee (CCITT) (1988). Recommendation X.509, The Directory - Authentication Framework, CCITT Blue Book, Vol III.8, 1988

Johnson, D.B. and Perkins, C. and Arkko, J. (2002) Mobility support in Mobile IPv6. IETF Internet Draft <draft-ietf-mobileip-ipv6-18.txt> (work in progress), June 2002

Lehti, I. and Nikander, P. (1998) Certifying Trust. In I Zheng, editor, Public Key Cryptography - First International Workshop on the Practice and Theory in Public Key Cryptography PCK'98, LNCS 1431, pages 83-98, Springer-Verlag, March 1998

Schulzrinne, H. and Casner, S. and Frederick, R. and Jacobson, V. (1996) RTP: A Transport Protocol for Real-Time Applications, Request for Comments 1889, 1996