إلى قارئ هذا الكتاب ، تحية طيبة وبعد ...

لقد أصبحنا نعيش في عالم يعج بالأبحاث والكتب والمعلومات، وأصبح العلم معياراً حقيقياً لتفاضل الأمم والدول والمؤسسات والأشخاص على حدٍّ سواء، وقد أمسى بدوره حلاً شبه وحيدٍ لأكثر مشاكل العالم حدة وخطورة، فالبيئة تبحث عن حلول، وصحة الإنسان تبحث عن حلول، والموارد التي تشكل حاجة أساسية للإنسان تبحث عن حلول كذلك، والطاقة والغذاء والماء جميعها تحديات يقف العلم في وجهها الآن ويحاول أن يجد الحلول لها. فأين نحن من هذا العلم ؟ وأين هو منا؟

نسعى في موقع عالم الإلكترون www.4electron.com لأن نوفر بين أيدي كل من حمل على عاتقه مسيرة درب تملؤه التحديات ما نستطيع من أدوات تساعده في هذا الدرب، من مواضيع علمية، ومراجع أجنبية بأحدث إصداراتها، وساحات لتبادل الآراء والأفكار العلمية والمرتبطة بحياتنا الهندسية، وشروحٍ لأهم برمجيات الحاسب التي تتداخل مع تطبيقات الحياة الأكاديمية والعملية، ولكننا نتوقع في نفس الوقت أن نجد بين الطلاب والمهندسين والباحثين من يسعى مثلنا لتحقيق النفع والفائدة للجميع، ويحلم أن يكون عضواً في مجتمعٍ يساهم بتحقيق بيئة خصبة للمواهب والإبداعات والتألق، فهل تحلم بذلك ؟

حاول أن تساهم بفكرة، بومضة من خواطر تفكيرك العلمي، بفائدة رأيتها في إحدى المواضيع العلمية، بجانب مضيء لمحته خلف ثنايا مفهوم هندسي ما. تأكد بأنك ستلتمس الفائدة في كل خطوة تخطوها، وترى غيرك يخطوها معك ...

أخي القارئ، نرجو أن يكون هذا الكتاب مقدمة لمشاركتك في عالمنا العلمي التعاوني، وسيكون موقعكم عالم الإلكترون ww.4electron.com بكل الإمكانيات المتوفرة لديه جاهزاً على الدوام لأن يحقق البيئة والواقع الذي يبحث عنه كل باحث أو طالب في علوم الهندسة، ويسعى فيه للإفادة كل ساعٍ ، فأهلاً وسهلاً بكم .

مع تحيات إدارة الموقع وفريق عمله

www.4electron.com

We Trip The Light
FANTASTIC

ELSEVIER

Computers
&
Security

www.elsevier.com/locate/cose

Number 1 February 2007

Contents

ELSEVIER

Computers & Security

From the Editor-in-Chief

What infosec changes are likely to result from the recent US election?☆

In the recent national election in the US the Democrats ended up with control of both branches of Congress. Widespread dissatisfaction with the war in Iraq coupled with concern over corruption by Republican legislators fuelled the Democrats' victory.

Many changes are likely to result from the recent election. The Democrats are, for example, now in a position to control and very possibly to limit funding for the Iraqi war. National economic and environmental issues are likely to rise to the forefront. Efforts to raise the national minimum wage are a certainty. The effect of the Democratic victory on infosec is, however, much less certain. Nevertheless, changes in two major areas, national legislation and the practice of infosec within the federal government, are likely to occur.

The first six years of total Republican control of the Executive and Legislative branches of the US government have been marked by little if any progress in passing computer crime-related legislation. Many computer crime-related bills have been considered in the House and in the Senate, but few of them have gone very far. The CAN-SPAM Act, a statute that represents only a minor step forward in the war against computer crime, is one of the few exceptions. Federal laws that prohibit the introduction of spyware into systems, that require suitable levels of protection for personal and financial information, regardless of whether it is stored or transmitted, and that require prompt notification of potentially affected individuals when data security breaches occur are most needed. The Republican Party's main agenda has been fighting terrorism; fighting computer crime has paled in importance. Additionally, this Party has the reputation of being pro-business. Legislation that requires businesses to implement additional security-related controls generally results in extra costs to businesses, something that has inhibited passage of security-related legislation that requires compliance in the business sector. Additionally, the Republican Party tends to eschew big government, i.e, government that continually interferes with and regulates organizations and individuals.

With the proverbial passing of the baton to the Democrats, more concerted efforts to pass national computer-crime related legislation are likely to occur. Additionally, legislation designed to protect individuals against identity theft and spyware is more likely to pass. Democratic Senator Diane Feinstein is, for example, more likely to be successful in her effort to pass legislation that requires notification of potentially affected individuals if data security breaches occur. At the same time, however, it is also important to not be too optimistic; the Democrats are in reality unlikely to get too far in their legislative efforts. The Democrats hold only a narrow lead in both branches of Congress, after all, and the Republican President has veto power over legislation. A two-thirds vote in the Senate is required to override a veto.

The second area to consider is the practice of infosec within federal departments and agencies. Whereas not much progress in federal anti-computer crime regulation occurred while the Republicans controlled the Legislative and Executive Branches of the government, the opposite is true as far as the practice of infosec within government circles goes. The OMB, GAO, and several Congressional committees exerted great amounts of pressure on departments and agencies to improve their security practices. This strategy was by all appearances mostly successful. Although government departments and agencies did not exactly serve as best practice models, they improved their security practices, as indicated by the generally higher marks that Congressional oversight committees gave them over time. Republicans also led the way in getting the Federal Information Systems Management Act (FISMA) passed.

The ''bottom line'' is that changes will almost certainly occur as the result of the Democrats' gaining control of Congress. Some of these changes are likely to be infosec related. A good start would be to pass legislation designed to protect individuals against identity theft, a risk that has been growing disproportionately over the years. Whether or not significant changes in infosec legislation and other areas will occur depends, however, on whether the Democrats can move forward despite the presence of numerous significant obstacles, one of the greatest of which is whether the Democrats can and will create a definitive agenda. And if and when such an agenda is created, the next question is whether or not it will include sufficient attention to infosec-related issues.

☆ The opinions in this editorial are entirely those of the author, not of High Tower Software. They do not in any way represent High Tower Software's position on the issues that are addressed.

Dr. E. Eugene Schultz, CISSP, CISM
High Tower Software
26970 Aliso Viejo Parkway, CA 92656, USA
E-mail address: eeschultz@sbcglobal.net

0167-4048/$ – see front matter

doi:10.1016/j.cose.2006.12.007

ELSEVIER

Computers & Security

Security views

1. Malware update

Up to 10,000 McDonalds customers in Japan received Flash MP3 players infected with a mutation of the QQpass spyware Trojan horse program. The MP3 players, given by McDonalds as prizes, were preloaded with 10 songs. The Trojan is capable of stealing passwords and other sensitive information. McDonalds has made an apology, set up an assistance line to help in finding and recalling the infected MPs players, and distributed procedures for deleting the QQpass Trojan.

A number of video iPods purchased after September 12 last year has been infected with the RavMovE.exe worm (which is also called the W32/Rjump worm). This worm infects Windows PCs and connected external drives when iPods are connected to infected computing systems. It also creates a backdoor on devices that it infects. Apple has not recalled infected iPods; up-to-date anti-virus signatures are effective in identifying and eradicating this worm. RavMovE.exe does not infect iPods or computers running Mac OS.

The SpamThru Trojan horse installs a pirated copy of anti-virus software on Windows systems that it infects. Once it has infected a system, it starts scanning the computer and erases any other malware during the next reboot in an attempt to monopolize computer resources. The Trojan, which is used to send spam for a "pump-and-dump" stock scam, uses peer-to-peer technology to communicate. If the control server is shut down, the individual who perpetrates a spam attack must merely control one peer to inform the others of the location of a new control server.

Some postings of the Google Video email group may have been infected by the W32/Kasper.A worm. The postings have apparently been removed, but Google still recommends that anyone who may possibly have an infected system due to interacting with this group run anti-virus software. Google has posted an apology on its Web site and has stated that it has implemented measures designed to preclude such infections from happening again.

The first instance of MacOS X spyware has surfaced. This proof-of-concept code could potentially be installed without users' awareness. The program, known as iAdware, installs itself as a System Library. It does not exploit any vulnerability per se, but instead takes advantage of a feature in Mac OS X that enables it to execute every time an application is loaded.

W32.Spybot.SCYR is spreading. It exploits six vulnerabilities, five of which are in Microsoft products and one of which is in Symantec's anti-virus tool. W32.Spybot.SCYR has infected numerous computers at universities in the US and Australia. Network connections to port 2967, which also occur when Symantec software is run, may indicate an infection by this Trojan. Patches for all of the vulnerabilities exploited by this piece of malware are available.

A few interesting new forms of malware have surfaced since the last issue of *Computers and Security*, yet once again nothing radically new in the malware arena has occurred. Although this would superficially lead one to conclude that viruses, worms and Trojan horse programs are becoming less of a problem, this unfortunately is not the case. Malware is still very much alive and well; it has simply become much more clandestine.

2. Update in the war against cybercrime

Four Russians have received prison sentences of eight years for their involvement in an extortion scheme. The perpetrators threatened to launch distributed denial-of-service (DDoS) attacks against on-line bookies and casinos in the UK if they did not pay the perpetrators a specific amount of money. Each of the recently sentenced individuals has also been fined 100,000 rubles. As many as nine persons may have participated in the extortion scheme, which crossed several countries' borders. The sentences for these crimes are the harshest ever for Russian computer crime.

Daewoo Hanel Electronic Corporation, an affiliate of Daewoo Corporation in Vietnam, must pay 15 million dong for using pirated copies of Microsoft Office and Windows, Auto CAD and other software. Vietnam's Ministry of Culture and Information announced that the pirated software was discovered during a recent unannounced inspection of the company's software. A Daewoo Hanel representative said that no one from this company knew that the software was illegal; it was pre-installed on systems that this company bought.

Parkev Krmoian of California has been charged with pilfering money from bank accounts of Dollar Tree customers in California and Oregon, allegedly using gift cards that were reprogrammed as ATM cards to steal money from these accounts. Law enforcement officers are also trying to find an additional person who appeared in surveillance photos taken at an ATM where the reprogrammed cards were used.

The US Securities and Exchange Commission (SEC) says that a dramatic increase in the number of attacks against on-line brokerage accounts such as Ameritrade and E-Trade accounts has been occurring. The individuals who are perpetrating these attacks are deploying keystroke loggers and spyware to illegally access accounts of unwary customers, steal money from these accounts, and/or to initiate unauthorized trades. Organized crime rings in Eastern Europe, particularly in the Ukraine, and Russia, appear to be responsible for these attacks. E-Trade losses in connection with these attacks have totaled more than USD 18 million during the last three months. Ameritrade says that it will compensate customers who lose money because of on-line fraud. Canada's Investment Dealers Association has been seeing similar attacks.

Illegal computer-related activity by high school students is increasing. Some North Branch, Minnesota high school students have been suspended for allegedly gaining unauthorized access to student and staff PIN numbers that are used in the school cafeteria and media center. Other student information was not illegally accessed, and no indications exist that any of the information that the accused students allegedly obtained was misused. The students are unlikely to face arrest. A computer lab manager for the high school found the security breach while cleaning up files. New PINs will be issued. This incident is similar to one in Janesville, Wisconsin, where a high school student allegedly gained unauthorized access to the school's computer system, causing troubles that resulted in a substantial loss of class and work time for the whole school district. The student was expelled. No evidence that any personal information has been compromised or that passwords, grades or student records were changed exists. Finally, a joint investigation by the Broward County, Florida School District and the Broward County Sheriff's Office is underway to determine if a student illegally accessed a Cooper City High School computer and altered grades. The possibility that attendance and community service records were also changed is also being investigated. Misuse of district technology constitutes a felony in Broward County; the perpetrator in this incident could thus face criminal charges.

Mathew Bunning, the former Australian drug squad detective, has received a jail sentence of nearly seven years for his having provided a drug dealer with information about police investigations in return for gifts. He ran a password cracker to obtain his colleagues' passwords, which he used to gain access to the information provided to the drug dealer. He had become morphine-addicted after a back injury.

Nine individuals in the Peoples Republic of China (PRC) have been sentenced to prison terms and must also pay fines of between 40,000 and 200,000 yuan for digital piracy-related activities. Four of those convicted received 13-year terms for creating and selling illegally copied materials, and another received a sentence of two years for selling bootleg software and DVDs.

Sweden's first music file-sharing-related trial and conviction ever occurred recently. A Swedish man whose identity has not been revealed was convicted for posting four copyrighted songs on the Internet and making them available for downloading. The International Federation of the Phonographic Industry (IFPI) claims the man had made 13,000 songs available for downloading, but the prosecution was able to obtain evidence concerning only four of the songs being shared. The man must pay a fine of 20,000 kronor. Sweden has recently tightened its music file-sharing statutes.

Matthew Decker of Kansas has received a sentence of five years in federal prison for gaining unauthorized access to US Army computing systems and pilfering information associated with between 250 and 300 credit card accounts and then using it to rack up USD 12,557 in unauthorized charges. Decker entered into a plea agreement in which he pleaded guilty to one count of illegally accessing a protected computing system and one count of possession of unauthorized credit card account access devices. Determining the damage resulting from the break-ins and restoring information and programs cost the US Army USD 25,000.

John Bombard of Florida has been charged with launching a distributed denial-of-service (DDoS) attack against caching service provider Akamai's Domain Name System (DNS) servers. He allegedly created a botnet by launching a variant of the Gaobot worm over two years ago; numerous Akamai client Web sites were knocked out of service. If convicted of the charges of deliberately gaining access to a protected computing system without authorization, he could be sentenced to up to two years of prison time and receive a maximum fine of USD 200,000.

Microsoft has won a violation of trademark case against an unidentified German spammer who sent unsolicited bulk email with falsified Hotmail return addresses without Microsoft's consent. The man must also pay all the costs in connection with the volumes of spam he sent.

Twenty-two persons in Finland must pay damages of more than EUR 420,000 for violating copyright laws by running a peer-to-peer file sharing network called "Finreactor." Among the numerous plaintiffs were software and media companies. Fines ranged between EUR 60 and 690 per person and investigation and court costs to be borne by the defendants amount to more than EUR 140,000.

Clarity1 Pty Ltd must pay a fine of AUD 4.5 million and Wayne Mansfield, the company's director, must pay AUD 1 million for sending 280 million unsolicited commercial email messages over a two-year time span. Additionally, Australia's Federal Court has prohibited Clarity1 from sending such messages in the future. The conviction was the first ever under Australia's Spam Act of 2003.

Terrence Chalk and his nephew, Damon Chalk, both of New York, face fraud and conspiracy charges. They are being accused of using the names, addresses and Social Security numbers (SSNs) of Compulinx employees for the purpose of obtaining loans, credit cards and lines of credit. Terrence Chalk owns Compulinx; he could face a maximum prison term of 165 years and a fine of USD 5.5 million. Damon Chalk could receive a prison sentence of up to 35 years and fine of USD 1.25 million.

The laptop computer of an employee of a Harrisburg, Pennsylvania water treatment facility was infected and then used to install malicious code on one of the facility's computer systems. Law enforcement said that the attackers did not target the facility per se, but instead intended to use the infected computer to spew email messages.

Four Chilean persons have been arrested on the grounds that they broke into NASA and Chilean finance ministry Web sites in addition to sites of other governments, including Israel, Venezuela, and Turkey. The number of compromised Web sites totals 8000. An eight-month long investigation in which Chilean police worked with Interpol and intelligence agencies in Israel, the US, and a number of Latin American countries.

A two-year long investigation involving the FBI and Polish law enforcement, "Operation Cardkeeper," has targeted an on-line black market for illegally obtained financial account information used in identity theft attempts. Fourteen persons, 11 of whom are Polish and three of whom are Americans, have been arrested so far, and more arrests of Americans and Romanians are probable. Eleven Polish and three Americans persons have been arrested; two more Americans are expected to be arrested soon.

A case brought against a Spanish man for illegal file-sharing was dismissed; the judge ruled that Spanish law does not prohibit downloading music for personal use. The individual in question allegedly downloaded songs and then offered them on CD through email and chat rooms. No evidence that the man profited from his alleged activity exists. Promusicae, Spain's recording industry federation, plans to appeal the ruling.

The US Federal Trade Commission (FTC) has levied a fine of USD 3 million on Zango (called "180Solutions" until recently). Zango was accused of downloading adware to computing systems in the US without users' consent and also of neglecting to make a way to remove the adware available. According to the FTC, Zango's programs were covertly downloaded more than 70 million times; more than 6.9 billion pop-up advertisements resulted. From now on, Zango will ask for consumers' consent before downloading programs to their computing systems and will supply a way of removing its adware. Odysseus Marketing and its chief executive, Walter Rines, have also consented to settle similar FTC charges that they broke federal law by distributing software that installs itself covertly on users' systems and then changes configuration settings. John Robert Martinson, principal of Mailwiper, Inc. and its successor company, Spy Deleter, Inc., will also settle FTC charges of downloading spyware onto users' computers and then bombarding them with adware that encouraged them to buy anti-spyware products. Finally, a US District Court judge in Nevada issued a temporary restraining order against ERG Ventures and one of its affiliates for allegedly surreptitiously installing spyware and other malicious code on users' computing system. An FTC complaint seeks a permanent restraining order against both of these companies on the grounds that they engaged in unfair and deceptive practices. Users downloaded free screensavers and video files, but unbeknownst to them a Trojan program, Media Motor, was downloaded to their computers. Media Motor then downloaded additional malicious programs.

Matthew Byrne, the UK man who gained unauthorized access to and defaced four profiles on the loveandfriends.com dating Web site and then attempted to extort money by threatening to erase the company's database, will not have to serve jail time. He instead received an eight-month jail sentence which was suspended for two years as well as two years of supervised parole. He pleaded guilty to unauthorized modification of a computing system in violation of section three of the UK's Computer Misuse Act (CMA). Extortion charges against him were also dropped. A law enforcement search of his home turned up evidence that he also wrote the Mirsa-A and Mirsa-B worms.

Adrian Ringland, a UK man, has been sentenced to 10 years of imprisonment for deceiving adolescent girls such that they downloaded a Trojan program that took control of their computers. Adrian Ringland also used pressure tactics to try to get the girls to send him nude photos of themselves. Once he had such photos, he tried to blackmail the girls into sending more photos. His arrest was the result of a joint investigation that included the UK Serious Organized Crime Agency, the Royal Canadian Mounted Police, the FBI, and Microsoft Corporation.

Garyl Tan Jia Luo, a 17-year-old polytechnic student in Singapore, faces up to three years of imprisonment and a maximum fine of SD 10,000 for allegedly "piggybacking" on a home wireless network. A neighbor of his filed a complaint against him that led to his arrest. He has been released on SD 6000 bail.

Joseph Harlen Shook of Florida has been indicted on the charges that he gained unauthorized access to a computer system of Muvico Theaters last year in May, resulting in disruption of the sale of on-line tickets and the processing of credit card transactions at six theaters. Shook was the director of information technology for the company until his position was eliminated shortly before the attack. If convicted of all the charges he faces, he could be sentenced to a maximum prison term of 10 years and a fine of USD 250,000. Investigators matched the ID of the device used to gain access to the Muvico Theaters system to a wireless adapter that Shook had. No customer information was accessed. He was released on USD 100,000 bail. The case is similar to one involving Stevan Hoffacker of New York, who faces one count of unauthorized access to a protected computer network for allegedly accessing his former employer's computer system without authorization. He is the former director of information technology and also vice president of technology at Source Media Inc. Hoffacker potentially faces up to five years of imprisonment.

The Seoul Metropolitan Police Agency's Cyber Terror Response Center arrested several phone sex company staff and one other person in connection with the compromise and misuse of customer information from rival companies' computing systems. The accused persons allegedly gleaned information pertaining to 8.5 million of their competitor's customers and then used it to send phone messages with sexual content, allegedly using phones registered in other names to send the messages. Additionally, cell phones were duplicated to circumvent charges for sending text messages.

Spanish law enforcement authorities have arrested two teenagers in Alicante for allegedly writing a Trojan program that they allegedly deployed to remotely control Web cams at a college and then allegedly used the embarrassing footage they obtained to blackmail those who were filmed. Two adults in Madrid who allegedly used a Trojan program that was based on the same kind of malicious code to pilfer information that they later used to commit credit card fraud were also arrested. The four arrests resulted from a Spanish law enforcement operation named "Operation Praxis."

Max Parsons of the UK was convicted of charges that he used his MP3 player to pilfer ATM customers' card information. He gleaned the information by plugging his MP3 player into free standing ATMs; he then used the stolen information to create bogus cards used to make purchases. Parsons received a prison sentence of 32 months.

According to L'Equipe, a daily sports publication in France, break-ins into computing systems at a French national anti-doping laboratory have occurred. The attackers reportedly accessed information and then sent the International Olympic Committee (IOC) and the World Anti-Doping Agency (WADA) letters written to cast doubt upon the laboratory's testing procedures by including information pilfered from the lab. The letters bore identification of a laboratory, Chatenay-Malabry. A possible suspect has been identified.

Microsoft has initiated legal action against 129 persons in Europe and the Middle East for their alleged participation in phishing schemes. Nearly half of the cases are based in Turkey. Microsoft's suits are in connection with its Global Phishing Enforcement Initiative launched last March. Settlements from the legal action range from fines of EUR 1000 to a 30-month jail sentence for a Turkish man.

The Software and Information Industry Alliance (SIIA) has reached a settlement for a case against two individuals who had been selling pirated copies of Norton security software on eBay for the last two years. Kevin Liu, GT Tian and Kevin Liu have agreed to pay USD 100,000 in damages and have consented to cease selling illegal software and to give SIIA their records of customers and suppliers.

The accelerated growth of computer-related crime continues, as once again shown by this long list of accounts of attempts to identify, charge, and convict perpetrators of this activity. I was, however, troubled by the case of Matthew Byrne. It is difficult to understand how in the light of his egregious deeds he was spared from having to serve even one day in jail. The judge in Byrne's case serves as an example of what is so often wrong with the legal system when it comes to dealing with computer crime. Worst of all, Byrne's having escaped significant punishment for his computer crimes will send a powerful message to the perpetrator community that punishment for computer crime is nothing to fear.

3. More compromises of personal and financial information occur

Many more compromises of personal and financial information have occurred. Computer theft and loss of computers proved once again to be one of the major reasons for such compromises, as per the following news items:

- A laptop computer on which personal information of 2400 residents of the Camp Pendleton Marine Corps base is stored is missing. Lincoln B.P. Management Inc., which manages housing on the base, reported that the computer was missing. Lincoln P.B. is informing people who have potentially been affected by this incident.
- Two computers pilfered from the house of a University of Texas at Arlington faculty member last year in September contain personally identifiable information pertaining to approximately 2500 students at this university. This information includes names, Social Security numbers (SSNs), grades and email addresses of students who enrolled in engineering and computer science classes between 2000 and 2006. A university spokesman said that potentially affected students are being informed of what happened. The university has posted a Web page containing pertinent information about the incident.
- A laptop computer holding SSNs of as many as 43,000 prior and current T-Mobile USA employees vanished after a T-Mobile employee checked the computer in at an airport. T-Mobile has sent letters to everyone whose data were stored on the missing computer and is offering them a year of credit monitoring at no cost to them.
- A laptop system taken from the car of an Allina Hospitals and Clinics nurse has information pertaining to approximately 14,000 individuals who have taken part in an obstetric home-care program since June 2005.
- A laptop system belonging to the art department at the University of Minnesota was stolen from a faculty member who was traveling in Spain. The laptop contains personally identifiable student data. This is the second recent laptop theft for this university; last year in September the university announced that two Institute of Technology laptops that held student data were stolen.
- A desktop system pilfered from Affiliated Computer Systems, which operates the Department of Human Services Family Registry, contains Child support payment-related information that includes personally identifiable information pertaining to many Colorado Department of Human Services clients. The computer was located in a physically secure area that was monitored by surveillance cameras. Potentially affected clients have been informed of the incident. Police detectives are cooperating with the Colorado Bureau of Investigation and Human Services Department officials in investigating the theft.
- A laptop system belonging to an insurance company in Plymouth Meeting, PA was stolen. On the computer were names, birthdates and driver's license numbers of over 1200 Villanova University students as well as staff who are insured to drive university-owned vehicles. Individuals whose information was on the stolen laptop were notified of the theft.
- Starbucks Corp. says that four of its laptop systems, two of which contain names, addresses and SSNs of about 60,000 current and prior employees, are missing. The computers disappeared from the Starbucks corporate support center in Seattle last September. The company is informing potentially affected individuals. No reports that the information has been misused have surfaced.
- The UK's Financial Services Authority (FSA) is investigating the theft of a laptop computer on which Nationwide Building Society customer information is stored. The computer was stolen from an employee's house last August. An FSA representative said that the information does not include PINs, passwords or information related to financial transactions, although exactly what information is on the stolen computer and how many people may be affected by this incident still remains unknown. Nationwide has begun informing its 11 million customers about what occurred.
- A laptop system on which personally identifiable information of Connors State College students and individuals who have been awarded Oklahoma Higher Learning Access Program scholarships is stored has been recovered. A

student at this college has been identified as a possible suspect in the theft.

- Two individuals have been arrested for their alleged involvement in the well publicized stealing of a laptop system belonging to the Transportation Department's Office of the Inspector General last summer in Miami. The laptop was pilfered from a locked car in a restaurant parking lot. Although the laptop has not been recovered, an investigation into the theft revealed the existence of a laptop theft ring in the area. Thieves appear to have motivated by the value of the laptops rather than for information stored on them.
- Three laptop systems pilfered from the offices of LogicaCMG contain sensitive financial information pertaining to more than 15,000 London Metropolitan Police officers. LogicaCMG provides outsourced management of payroll and pension payments. A man has been arrested for his alleged involvement with the theft.
- A laptop system stolen from the Ontario Science Centre has a database with members' registration information – names, addresses and credit card information. Access to the laptop and the database requires entry of separate passwords. The laptop was stolen from a locked office last September. The Ontario Science Centre has informed potentially affected members by postal mail. An investigation is being conducted.
- Two computing systems pilfered from a Jeffersonville, Indiana health center last November contain personal information pertaining to more than 7500 Indiana women. The health center was under contract with the state of Indiana to manage information for the state's Breast and Cervical Cancer Program. The information stored on the computers includes names, addresses, dates of birth, SSNs, and medical and billing information. Access to the information is password protected at two different levels. The health center sent letters to the women who were potentially affected to inform them of the theft.
- Kaiser Permanente has announced that a laptop system stolen from an employee's locked car in California holds sensitive medical information for nearly 40,000 of its Denver area patients. All information was password-protected, and some of it was encrypted. Kaiser Permanente informed potentially affected patients of the incident.

Other compromises resulted from unauthorized access to systems, as described below.

- Personal information pertaining to UK citizens may be being stolen from India call centers and then sold to the highest bidder. The information may include credit card information, passport and driver's license numbers, and bank account information. The perpetrators may also have access to taped conversations with US consumers in which personal and financial information such as credit card numbers is exchanged.
- The University of Iowa has contacted 14,500 persons whose SSNs were on a computing system to which an attacker gained unauthorized access. The people had participated in research studies concerning maternal and child health for over ten years. The attacks were automated and appeared to be motivated by the desire to locate places in which to store digital video files. No evidence that the SSNs and other personal information were accessed exists. Law enforcement has been informed about the security breach. The University of Iowa has set up an FAQ web page to provide information about the incident and to answer questions.
- A perpetrator gained unauthorized access into a Brock University computing system and accessed personal information pertaining to approximately 70,000 university donors. The fact that the perpetrator logged in to the system right away indicates that the perpetrator already knew the password. The compromised information includes names, addresses, email addresses and in certain cases, credit card and bank account information. Individuals whose credit card and bank account information was exposed received phone calls within a day of the university having learned of the data security breach. The others who were potentially affected were mailed letters that informed them of the incident.
- The Congressional Budget Office (CBO) has stated that attackers gained unauthorized access to one of its servers and stole email addresses of mailing list subscribers. The vulnerability that the attacker exploited has been fixed, but since the intrusion the attackers have mailed phishing messages that appear to originate from CBO to the stolen addresses. Law enforcement has been informed of the security breach and has started to investigate.
- Perpetrators gained unauthorized access to two US Virgin Island government accounts at Banco Popular and pilfered USD 500,000 from the accounts. The bank has restored USD 300,000 to the accounts; the remainder of the money is likely to be returned soon. The perpetrators stole the money little-by-little over two months.
- A break-in into two computer databases at Children's Hospital in Akron, Ohio has exposed personal data pertaining to approximately 230,000 patients and family members and 12,000 financial donors. Although aware of the security breach soon after it occurred, hospital officials did not contact law enforcement until weeks later. Consultants had at first told these officials that the incident was not serious, but over time these officials became aware the incident was more serious than they initially believed. The hospital has mailed letters to inform those potentially affected by the incident.

Other personal data exposure incidents were the result of inadequate protection of personal information on Web sites:

- A Florida woman reportedly learned that her marriage license could be seen on the Orange County, Florida controller's Web site after someone filled out a loan application using her name. Information in her marriage license included her and her husband's name, date of birth and SSN. Orange County officials are reportedly paying a vendor USD 500,000 to black out every SSN on the Web site by January next year.
- The Bowling Green Ohio police department posted the incorrect version of a report on its police blotter Web site. Posted reports on this site usually have personally identifiable

information removed, but the incorrect report version, called an "end of day report," revealed the birth dates, SSNs, driver's license numbers and other data pertaining to every individual with which Bowling Green police came in contact that day. The exposed information is no longer on the site and a cached version of the report was removed from Google servers.

Several data security breaches were due to missing or stolen media:

- A Port of Seattle spokesperson announced that six disks have disappeared from the ID Badging Office at the Seattle-Tacoma International Airport (SEATAC). The disks hold sensitive personal information of nearly 7000 current and previous SEATAC employees such as names, SSNs and driver's license numbers scanned from paper forms. Individuals who are potentially affected by this incident will be informed by Postal Service mail; these people will be identified because the data were backed up.
- The Sisters of St. Francis Health Services are mailing letters to over 250,000 patients whose personal information that included patient names and SSNs was on CDs that could not be found for a while. A contractor had copied information from hospital files to the CDs to do work at home; the CDs were in a computer bag that someone returned to a store to get a refund. An individual who purchased the returned bag found and returned the CDs. The incident potentially affects patients from 12 hospitals, 10 in Indiana and 2 in Illinois. St. Francis did not notify potentially affected individuals until approximately two months after the incident occurred.
- A hard drive containing the names of SSNs of 400 or more air controllers is missing from the Cleveland Air Route Traffic Control Center in Oberlin. A Federal Aviation Administration (FAA) spokesperson says that the agency thinks that the drive was encrypted; the FAA is investigating what happened to decide whether the drive was actually stolen.
- A thumb drive on which names, SSNs, and other personal information of current and prior employees at the Portland, Oregon International Airport is missing from the Transportation Security Administration's (TSA) command center there. The federal security director at this airport says the drive was in all likelihood accidentally thrown out.
- A disk belonging to KSL Services and that contained personally identifiable information of approximately 1000 Los Alamos National Laboratory (LANL) contract workers has disappeared. On the disk are data pertaining to KSL employees; LANL-related information is not on this disk.
- Approximately 200 pages of classified documents and a USB drive storing classified material were found at the home of another former LANL contractor. The classified documents were mostly older ones that are no longer considered important. LANL officials conceded, however, that some of the documents were moderately important. The FBI has launched an investigation.

Some additional news items concerning data security breaches include:

- Scotland Yard is investigating how credit card information and passwords from thousands of PCs in the United Kingdom and possibly tens of thousands more in other countries have been stolen. The stolen information was found on computers in the US. Law enforcement is informing individuals whose information was stolen.
- A laptop system that was once owned by Intermountain Healthcare in Utah supposedly had its hard drive scrubbed before it was donated to Deseret Industries. The person who purchased the laptop, however, found a file on the hard drive of the donated computer that held personally identifiable information that included names and SSNs of more than 6000 individuals who had been employed by Intermountain Healthcare between 1999 and 2000. Potentially affected persons have been informed of the data exposure. Intermountain stopped using SSNs as unique employee identifiers several years ago, and now destroys hard drives when they are no longer being used.
- A computer belonging to Hertz car rental on which the names and SSNs of most US Hertz employees were stored was found at the home of a prior employee of this company. Law enforcement is investigating. Hertz announced that all employees whose information was on the computer will be informed of the incident. The former employee was allowed to access this information in connection with job-related duties.

I am continually amazed by the sheer number of data security breaches that occur. One would think that by now organizations would realize that determined perpetrators are continually trying to gain unauthorized access to personal and financial information, and that a large proportion of the perpetrators is anything but amateurs. Clearly, what we are seeing here is a lack of due diligence on the part of a high percentage of these organizations. Unfortunately for them, it will in most cases take a widely publicized data security breach or a lawsuit or two by angry individuals whose data have been compromised to bring these organizations out of their catatonic state when it comes to data security practices.

4. Number of pieces of compromised personal data approaches 100 million

The Privacy Rights Clearinghouse recently reported its tally of the pieces of personal and/or financial information that have been compromised in data security breaches. According to this organization, almost 94 million instances of such information being exposed have occurred since February 2005 when it began tallying this information. The updated tally includes thousands of recent data security breaches, including 9250 customer credit card numbers lost by apparel retailer Life is Good and large numbers of student records illegally accessed at US colleges and universities. Many of the reasons for data thefts and losses have little or nothing to do with technology; human error is far more likely to be a cause than intrusions or computer malfunctions. Data security exposures include loss of USB drives, laptops being stolen or misplaced, accidental printing and distribution of customer names, credit card and/or account numbers, and backup tapes being lost while in transit to storage facilities. Since 2001, approximately 1100 laptop computers belonging to the

US Commerce Department are missing. Last August, AT&T disclosed that credit card numbers and other personal data pertaining to 19,000 customers who had purchased DSL equipment from its on-line store were compromised when perpetrators were able to access a database. Last September, a third-party credit card service for Circuit City accidentally discarded five backup tapes containing information pertaining to 2.6 million prior and current Circuit City credit card holders.

The validity of statistics in the information security arena is often (if not usually) somewhere between somewhat and completely questionable. I nevertheless included this news item because it provides at least some indication of the magnitude of the data security breach problem that has been occurring over the years. At the same time, however, with the exception of Europe, little is being done to remedy the problem. The question thus becomes how great the toll in terms of the amount of exposed personal and/or financial information will have to be before the tide turns.

5. Telecom and Internet service providers to publicize security lapses

European Commission legislation due to go into law late this year will force telecom providers and Internet service providers to disclose if customer data are at risk. The Review of European Union (EU) Regulatory Framework for Electronic Communications Networks and Services mandates that all suppliers of electronic communications networks or services inform regulators and customers of any security breach that exposes customers' personal information to unauthorized access. The earlier Data Protection Act states that appropriate technical and organizational measures must be followed to thwart unlawful access to personal information. What is different about the new legislation is that it requires that the public be informed of data security breaches. The Data Protection Directive required telecommunications providers to retain sensitive information longer. If more information is retained longer, however, the likelihood of the information falling into the wrong hands also increases.

One thing that has puzzled me when I have been assembling news items to be included in the section on personal and financial information compromises is why news items concerning data security breaches in countries other than the US and Canada are not available. Are far fewer data security breaches occurring in other countries? After talking to numerous European colleagues and friends, I am not confident that this is true – many of them have assured me that there are also more than a few data security breaches in European countries, but that these breaches do not get reported and publicized the way that they do in the US and Canada. The new European Commission legislation requiring disclosure to the public if data security breaches occur is likely to have an enormous effect on the disclosure of such incidents. It is thus reasonable to expect that in time more news items about data security breaches in Europe will appear.

6. Spamhaus now fighting US Court judgment

Spamhaus will appeal a US court's USD 11.7 million judgment against it for putting e360 Insight on its spam blacklist. Spamhaus had originally indicated that it would disregard the court's ruling on the basis that US courts do not have jurisdiction over organizations in other countries such as the UK. e360 Insight then requested that the spamhaus.com domain be suspended, something that prompted Spamhaus to appeal the ruling against it rather than to ignore it. A judge, however, rejected e360's request, saying that it would be unduly punitive because all legal activities of Spamhaus would also be stopped. Ordering ICANN to suspend the domain would also, according to the judge, be inappropriate given that ICANN had nothing to do with Spamhaus' having put e360 on its spam blacklist. Additionally, ICANN said that it does not have the authority to suspend Spamhaus' domain. Meanwhile, Spamhaus continues to maintain that e360 Insight engages in sending spam.

This is an extremely intriguing case, one that is likely to set a precedent for future cases in which computing-related "good Samaritan" organizations take action against an organization or individual in the name of benefiting the public. If the organization or individual sues the good Samaritan organization in a country in which the good Samaritan organization is not based and wins, a game of chess ensues. There currently appears to be a stalemate in the Spamhaus case, but I seriously doubt that e360 has made its last move. The prospect of being awarded nearly USD 12 million and also having its name cleared because of a US court-mandated apology by Spamhaus is compelling motivation for e360 to do whatever is necessary to win. At the same time, however, it is difficult to envision what additional moves e360 can and will make until Spamhaus' appeal is heard. One thing is sure – computer crime cases and cases in which network and computing restrictions are placed against an entity that is ostensibly acting inappropriately too often produce inconsistent court rulings. What is deemed illegal or inappropriate at one level of the legal system may be ruled perfectly legal or appropriate at a higher level. It will thus also be extremely interesting to learn the outcome of Spamhaus' appeal.

7. SWIFT under fire for providing customer transaction information to the US

SWIFT (The Society for Worldwide Interbank Financial Telecommunications) has recently come under fire by several entities within Europe for providing customer financial transaction-related information to the US. The first such entity was the Belgian Data Privacy Commission, which asserted that SWIFT broke privacy laws when it provided information to various US agencies, including the CIA, to allow them to identify possible funding for terrorist activity by examining the many financial transactions that SWIFT processes. More recently, a European Union panel charged with monitoring SWIFT-related legal and privacy matters ruled that SWIFT violated EU data protection rules when it provided information concerning customer financial transactions to the US. The panel also asserted that SWIFT customers (i.e., financial institutions) share the blame with SWIFT for the violation of European privacy statutes that it says has occurred. The panel recommends that SWIFT either cease violating the statutes or face sanctions. Additionally, the EU will decide whether or not to take Belgium to court for not forcing SWIFT, which is Belgium-based, to confirm to EU data protection statutes.

Predictably, the US has defended the program in which it has received information about customer financial transactions from SWIFT as part of an anti-terrorist effort that began shortly after the September 11, 2001 terrorist attacks in the US. European critics maintain, however, that this program sacrifices Europeans' civil rights in the name of US security interests. SWIFT, on the other hand, claims that it has provided only a limited amount of information on financial transactions and has placed exceptional levels of protections and controls on the data. SWIFT also maintains that cooperating with the US in this manner is legal and is also essential in the fight against terrorism.

The current controversy concerning whether or not SWIFT should be turning customer financial transaction data over to the US is really nothing more than a continuation of the ongoing controversies growing out of the fact that different privacy protection standards in Europe and in the US exist. I cannot blame Europeans in the least bit for being upset over finding out that SWIFT has been providing customer data to the US. To say that the US has inadequate privacy protection requirements is a major understatement. Europeans thus have a very legitimate complaint. If the US needs data such as SWIFT financial transaction data as badly as the Bush Administration says it does, it would seem logical that the US should be willing to make some concessions, such as tightening privacy protections to the level required by EU privacy statutes. Unfortunately, however, once again this administration appears unwilling to try diplomacy and cooperation in dealing with international problems.

8. E-voting problems persist

Using only publicly available technical reference materials from the Irish Department of the Environment, Dutch IT specialists successfully used simple radio receivers to alter the functionality of Nedap e-voting machines recently acquired by The Netherlands and similar to the ones acquired by Ireland. The compromised machines were then reprogrammed to record inaccurate voting preferences and to play chess. According to the anti e-voting group, ''Wij vertrouwen stemcomputers niet'' (We don't trust voting computers), that organized the security test, a small part no larger than a postage stamp was replaced to compromise the machines. A similar technique could thus also be used on the Irish e-voting machines. This strengthened the position of the Dutch group and the Irish Citizens for Trustworthy E-Voting (ICTE), a lobby group, that no voting system should be used without a voter-verified audit trail. Furthermore, these groups point out that not only is the software vulnerable, but the hardware itself also poses a serious risk.

Dutch intelligence service AIVD determined 1200 Sdu e-voting computers were inadequate for the Netherlands' 2006 national elections after testing showed the machines could be easily accessed and controlled from 20 to 30 m away. Voters in many large Dutch cities had to cast their ballots with pencil and paper, although some municipalities could still decide to use different polling computers. The remaining 90% of the Netherland's polling computers are made by Nedap, which says it is installing new microchips and software to preclude voting manipulation.

The FBI is investigating the possible theft of software for electronic voting equipment on disks that were delivered anonymously to a former Maryland legislator. Three computer disks that reportedly contained key portions of programs used in the Maryland 2004 elections and created by Diebold Election Systems were accompanied by an unsigned letter critical of Maryland State Board of Elections Administrator Linda Lamone. On the disks were the logos of Cyber, Inc. and Wyle Laboratories, two testing companies that sent similar disks to the Maryland board after conducting tests on the Diebold equipment. Also on the disks were labels that indicated they contained source code for Diebold's Ballot Station and Global Election Management System (GEMS), a tabulation program used to count votes, programs. The release of the software could further expose potential security vulnerabilities in e-voting systems.

Informing the state board of elections that it was doing a ''technical refresher'' of installed e-voting machines, Diebold Election Systems covertly replaced defective motherboards in 4700 Maryland e-voting machines in Dorchester Allegany, Prince George's, and Montgomery counties in 2005 to fix screen freezes the company had discovered three years earlier. Although such freezes do not cause votes to be lost, they could confuse voters and election judges who might wonder if all votes on a machine might not be counted. The freezes were unrelated to issues raised in every Maryland precinct during last September's primary election; voter registration machines rebooted then without any warning. The cause of the rebooting was a software bug.

In a ruling contesting e-voting, a Colorado State District Court judge determined that the state's testing and certification process for their touch-screen e-voting machines did not use acceptable sets of standards, failed to include adequate testing procedures, and lacked adequate security methods. However, the judge did not ban the use of the machines during last November's election because decertifying them so close to the election would cause more problems. The state certification process did not cover all potential threats to the machines and the machines had not been tested for vulnerability to malicious code access. The judge ordered that more comprehensive certification and security measures be developed in time for future elections.

A display problem in Virginia e-voting machines that increased the font size and thus distorted some of the names of some Virginia candidates on the candidate summary page once again illustrates the many problems e-voting machines are causing. Virginia election officials stated that the pages on which voters make their selections contain the full names – the summary page shows the voters all the possible selections before the vote is cast. The officials say they will have the problem corrected by the time of the elections later this year. Virginia lawmakers introduced a bill in the General Assembly that would require the State Board of Elections to design a pilot program to test e-voting equipment and paper records. The bill was not approved, however.

Security and reliability problems continue to surround the e-voting machine arena. Demonstrations seem to repeatedly show how the functionality of these machines can be modified without authorization. Reliability problems are also significant because they go hand-in-hand with security problems. If voting machines do not function reliably during elections, many security exposures that could allow dishonest persons to tamper with vote tallies also exist. As I have said many times before, the integrity of voting results is critical in democracies. If e-voting machines are used, there should be overwhelming assurance that they are 100% correct in tallying votes. The current state of the art in e-voting technology right now cannot provide this level of assurance, however; as such, e-voting needs to be kept away from elections until it can.

9. Business Software Association of Australia Boosts bounty for turning in users of pirated software

Realizing that the 31% software piracy rate in Australia is excessive by international standards, the Business Software Association of Australia boosted the reward for turning in people who use pirated or otherwise illegitimate software applications to AD 10,000, double the previous amount. To help make certain that any claims concerning possession of illegal software are bona fide, informants will be required to sign affidavits and also to participate in the legal process, the latter of which often takes considerable time. No criminal case of software piracy in Australia has occurred so far, possibly because the Australian Federal Police does not have enough resources to enforce Australia's copyright laws.

From what I have read, former employees of organizations that use illegal software are the most likely to report software piracy. Presumably, then, these employees are motivated more by revenge than anything else. I would thus question whether doubling the bounty for turning in people who use pirated software will do all that much good. On the other hand, if no other strategy is working all that well, this new strategy may very well be worth a try.

10. Researchers present new stealth encryption method

At the recent yearly meeting of the Optical Society of America, Princeton researchers Evgenii Narimanov and Bernard Qiang Wu presented a paper introducing a stealth messaging method that conceals messages within the noise of a fiber-optic transmission line infrastructure. This method can now conceivably be used to enable governments and businesses a method to provide inexpensive, widespread, and secure transmission of confidential and sensitive information. The sender must first translate a message into an ultrashort pulse of light. A commercially available Code Division Multiple Access (CDMA) encoder next spreads the intense, short pulse into a long, faint stream of optical information that is weaker than the noisy jitters in fiber-optic networks. The intended receiver decodes the message using information concerning how the secret message was originally spread out and using another CDMA device to bring the message back to its original state of compression. The researchers assert that this method is extremely secure. Even if eavesdroppers knew that a secret transmission was being sent, any incorrect knowledge of how the secret signal was diffused would make picking out the signal from the more intense public signal too difficult.

This is another intriguing new development in that it suggests a new and powerful method of secure message transmission that is quite different from conventional cryptography. The goodness of methods such as the one that Narimanov and Wu have described depends on the work effort to break them. So far, no one has attempted to break this new method, so we cannot really appreciate how strong this method is. Still, the fact that a new, quite different way of protecting transmitted message content has surfaced is significant in and of itself.

11. IFIP targets peer-to-peer file sharers again

Another crackdown on peer-to-peer file sharers is underway. The International Federation of the Phonographic Industry (IFPI) is focusing its efforts on people who put their music files onto peer-to-peer file sharing networks. More than 8000 people who have allegedly participated in peer-to-peer music sharing are now confronted with legal action. The new cases cover file sharers in 17 different countries who have been using BitTorrent, eDonkey, SoulSeek, WinMX, and other sites. Legal action has also been expanded to Poland, Brazil, and Mexico. The IFPI stated that a wide-variety of individuals – even parents whose children have engaged in peer-to-peer file sharing – has been targeted in the recent crackdown. Although recording industry representatives assert that this industry has been adversely affected by illegal downloading, skeptics counter that the IFPI is targeting users who buy most of their music and that declining CD sales are due to a variety of other factors.

The IFIP must have coined the phrase: "If at first you do not succeed, try, try again." Previous attempts to stem the tide of illegal music copying and sharing have been at best only slightly successful, so IFIP has initiated yet another similar strategy. One must at the same time, however, wonder why if one crackdown after another has not been all that successful in the past, IFIP now expects yet another crackdown to somehow be successful. There are many possible approaches to countering music piracy; my suspicion is that IFIP has just not yet thought of and tried the right one.

12. FBI asks ISPs to capture users' activities

Because of the need to monitor on-line criminal activity, FBI Director Robert Mueller asked that Internet Service Providers (ISPs) capture the on-line activities of their customers. This request dovetails with other law enforcement agencies' assertions that by the time they make contact with ISPs, customers' records may already have been erased. One proposal would mandate that registrars of domain names also maintain records. In private meetings with industry officials, Justice Department and FBI representatives, the topic of conversation turned to forcing search engine providers such as Google to keep their logs. The 1996 Federal Electronic Communication Transactional Records Act regulates data preservation and requires ISPs to retain any record in their possession for 90 days if a governmental entity requests that they do so. An additional federal law mandates that ISPs report evidence of child pornography to the National Center for Missing and Exploited Children, which must then by law turn in a report of child pornography activity to the appropriate law enforcement agency. When formulating its data retention rules, the European Parliament approved a requirement that communications providers within its member countries, several of which have already passed their own data retention laws, must keep customer-related information for a minimum period of six months and a maximum of two years. This requirement goes into effect next year. The European requirement applies to a wide range of traffic and location information that includes the identities of customers' correspondents, the date, time, and length of phone calls, VoIP (the voice over Internet Protocol) calls or email messages, and the physical location of the device used in the communications. The content of the communications does not need to be preserved, however.

I fear that ISPs and search engine providers are fighting a losing battle in their struggle to keep from being forced to hand over records of customers' on-line activities to law enforcement. As in the case of FBI Director Mueller, law enforcement's initial approach has been to ask for these records. ISPs and search engine providers have not complied with these requests in the past and are unlikely to do so both now and in the future, forcing law enforcement to go to court to obtain the information it needs. Law enforcement will argue that it needs this information to combat terrorism, whereas ISPs and search engine providers will argue that turning over records of customer activities would constitute an infringement of privacy. Given that privacy protection is almost non-existent in the US, it is safe to say that the courts will side with law enforcement.

13. Exposures in use of RFID technology found

Using readily available and inexpensive off-the-shelf radio and card reader equipment, University of Massachusetts security researchers Kevin Fu and Tom Heydt-Benjamin demonstrated how easy it is to glean sensitive personal information off of Radio Frequency Identification- (RFID) based credit and debit cards without having to make physical contact with the cards. Although their demonstration did not result in the ability to harvest verification codes normally needed in purchasing goods and services on-line, it is still possible for perpetrators to use the information that can be readily gleaned to order goods and services from on-line stores that do not require this information. Although card-issuing companies assert that data contained on RFID-based credit cards would be encrypted, Fu and Heydt-Benjamin, who are also members of the RFID Consortium for Security and Privacy (RFID-CUSP), found that the majority of the cards they tested used neither encryption nor other data security techniques. Card issuers have recently been modifying their cards so that the names of card holders are not sent to card readers.

The UK has issued more than three million hi-tech passports designed to frustrate terrorists and fraud perpetrators. The US informed 27 countries that participated in a visa waiver program that travelers to the US from those countries with passports issued after 26 November last year would either have to have micro-chipped biometric passports or would have to apply for a US visa. These biometric passports incorporate RFID chips containing the holder's personal data and a digital representation of their physical features. The new passports are protected by 3DES encryption. The International Civil Aviation Organization (ICAO), which develops standards for the passports, has recommended that passports contain biometrics based on facial scans, although countries could introduce fingerprints later. Unfortunately, the ICAO recommendations also included a faulty method for key creation. The key needed to access the data on the chips would be the passport number (first), the holder's date of birth (second) and the passport expiry date (third), all of which are contained in clear text on the printed page of the passport on a machine-readable zone. When the passport is swiped through a reader, all information needed to create the key is available; the microchip reader can then communicate with the RFID chip. The data, including the holder's picture, are then displayed on the official's screen. Because non-secret information printed in each passport is used to create each key, the key can be broken through cryptanalysis without a massive amount of effort. Additionally, a UK computer specialist used an easily and cheaply obtained reader to access the data on three of the new passports and was able to download the data to his laptop system.

RFID chip technology is regarded by many in the credit card industry as well as in government circles as the technology of the future. Although this technology offers several important advantages, namely ease of use, speed and reliability, its potential for security does not presently appear to be one of them. The good news is that a few modifications such as preventing cardholder names from being transmitted to card readers and using better methods of key generation may prove to serve as suitable countermeasures to reduce the potential for exploiting features of this technology to perpetrate identity theft and other types of crime.

14. New UK computer crime-related legislation passes

To close numerous loopholes in previous legislation, a law was passed in the UK to make denial of service (DoS) attacks punishable by up to 10 years in prison. The UK's Computer Misuse Act (CMA), passed in 1990, describes an offense as doing anything with criminal intent that "causes an unauthorized modification of the contents of any computer." As originally worded, unfortunately, the CMA did not contain specific language to cover DoS attacks. A little over a year ago a court cleared teenager David Lennon of charges of sending five million emails to his former employer because the judge decided that no offense had been committed under the provisions of this Act. Arguing that the company's server's purpose was to receive emails, and therefore the company had consented to the receipt of emails and their consequent modifications in data, Lennon's attorneys won. The judge concluded that sending emails is an authorized act and that Lennon had no case to answer – no trial took place. That ruling was later overturned, however, and Lennon was sentenced to two months' curfew with an electronic tag. Amendments to the 1990 legislation have already been included in the Police and Justice Bill. The Police and Justice Act of 2006 increases the penalty for unauthorized access to computer material from a maximum of six months' imprisonment to two years and expands the 1990 Act's provisions concerning unauthorized modification of computer material to make illegal having the necessary intent and requisite knowledge to maliciously impair the operation of any computer, to stop or hamper access to any program or data held in any computer, or to harm the operation of any program or information held in any computer. The targeted computing system, program or information does not need to be specified. The wording in the Act is sufficiently broad that paying someone to instigate an attack would still be defined as a crime with a penalty of up to 10 years of imprisonment. Furnishing the software to instigate the attack or offering access to a botnet is punishable by up to two years of imprisonment.

A new anti-fraud bill has been passed into law in England and Wales to close numerous loopholes in prior anti-fraud legislation that was written before newer methods of on-line criminal activity emerged. The Fraud Act 2006 received Royal Assent last week and goes into effect early this year. Until now there has been no single, general fraud law in English law. Scotland has a common law crime of fraud; fraud is committed when someone obtains a practical result using a phony pretense. The Fraud Act introduces a general offense of fraud that can be committed by false representation, by failing to disclose information, or by abuse of position. This Act also stipulates that writing software with the intent of using it in fraud is punishable by up to 10 years of imprisonment. The new law should help counter identity theft, spoofing, phishing, and possibly other types of attacks.

Updating the CMA was well overdue. This CMA was passed when computer and network technology was far less sophisticated than now and also when computer criminals' favorite type of attack was brute force password entry. Much has changed since then. According to CERT/CC, DoS attacks are now the most common type of Internet attack. Defining DoS attacks as a punishable crime is just one of the many improvements in this new legislation. The newly passed anti-fraud bill in England and Wales also substantially expands the definition of computer crime as well as other types of crime. Both pieces of legislation will help considerably in the war against computer crime in the UK, and will also serve as models of the types of computer crime legislation that are so badly needed in many other countries.

ELSEVIER

Computers & Security

Biometric attack vectors and defences

Chris Roberts

Department of Information Sciences, Otago University, Dunedin, New Zealand

Keywords:
Biometric
Identification
Security
Attack vector
Threat
Countermeasures
Defences

ABSTRACT

Much has been reported on attempts to fool biometric sensors with false fingerprints, facial overlays and a myriad of other spoofing approaches. Other attack vectors on biometric systems have, however, had less prominence. This paper seeks to present a broader and more practical view of biometric system attack vectors, placing them in the context of a risk-based systems approach to security and outlining defences.

1. Introduction

1.1. *Structure of this paper*

This paper contains the following:

- an introduction to the topic of biometric attack vectors;
- a brief review of previous models and a suggested new approach;
- an outline of the risk context; and
- a description of defences and countermeasures.

1.2. *Definitions*

For the purposes of this paper an *attack vector* is defined as the channel, mechanism or path used by an attacker to conduct an attack or to attempt to circumvent system controls. A *threat* is the possibility of an attack. *Spoofing* is the presentation of an artefact, false data or a false biometric claiming to be legitimate, in an attempt to circumvent the biometric system controls. A system *vulnerability* is a design flaw or feature that creates a security weakness and presents an opportunity for attack or exploitation of the biometric system.

1.3. *Problem outline*

The majority of reported biometric systems' incidents are related to spoofing. While some attempts have been made to represent a more complete view of attack vectors, successive representational models have become increasingly complex with decreasing practical application. Practitioners and information security professionals will seek structured and practical representations that correlate with existing methods and approaches to risk and security management. This paper presents such an approach.

1.4. *Preamble*

Biometrics are increasingly being used for security and authentication purposes and this has generated considerable interest from many parts of the information technology community. There has also been a great deal of interest from those interested in examining and researching methods of circumventing and compromising biometric systems.

In common with all security systems, there have been attempts to circumvent biometric security since they were introduced. Designing secure systems can be challenging and it is important to assess the performance and security

E-mail address: makiwa@paradise.net.nz

of any biometric system in order to identify and protect against threats, attacks and exploitable vulnerabilities. Security breaches are, most commonly, the result of an exploited vulnerability (Ratha et al., 2001). This includes poor physical security which continues to be an easily exploitable attack vector.

Often these vulnerabilities were not considered or had been discounted as implausible in systems design and management. It is, therefore, important to adopt a systems approach and assess *all* risks as failing to assess any one aspect can lead to a catastrophic failure of system security.

1.5. *Biometric spoofing history*

An early report into fingerprint devices and their susceptibility to acceptance of "lifted" fingerprints or fake fingers, was published by Network Computing in 1998 (Wills and Lees, 2006). They found that four out of six devices tested were susceptible to fake finger attacks.

Further research was undertaken by Tsutomu Matsumoto who published a paper on "gummy" fingers in 2002 (Matsumoto et al., 2002). In this research, finger sleeves were made from gelatine, designed to cover a fingertip and with a fingerprint on the outer surface. In testing, these had a high acceptance rate from fingerprint readers using optical or capacitive sensors. In addition, fake fingers could be enrolled in the system (68–100% acceptance).

In November 2002 c't magazine (Check et al.,) published the results of the testing of a variety of biometric devices. A number of spoofing attacks were successful, as were "man-in-the-middle" attacks on datastreams. Tests were conducted on fingerprint, facial recognition and iris scan biometric devices. The facial recognition devices were spoofed by playing back a video of a person's face. Iris scanners were spoofed with a high resolution photograph of an iris held over a person's face and with a hole cut in the photograph to reveal a live pupil. Another method of spoofing iris scanners is to replay a high resolution digital image of the iris.

In August 2003, two German hackers claimed to have developed a technique using latent prints on the scanner and converting them to a latex fingerprint replacement, small enough to escape all but the most intense scrutiny (Harrison, 2003). This method uses graphite powder and tape to recover latent prints which are digitally photographed, and the image enhanced using graphics software. Where complete fingerprints are not available, the graphics software is used to compile a fingerprint from overlapping portions recovered from the scanner.

The image is photo-etched to produce a three-dimensional reproduction of the fingerprint. This etch is then used to as a mould for the latex fingerprint.

More recently (December 2005), research undertaken at Clarkson University revealed that it was possible to demonstrate a 90% false verification rate in the laboratory (Clarkson University Engineer, 2005). This included testing with digits from cadavers, fake plastic fingers, gelatine and modelling compounds. However, when "liveness" detection was integrated into the fingerprint readers, the false verification rate fell to less than 10% of the spoofed samples.

Much of the activity in spoofing biometric systems has, up until now, been confined to researchers. However, as biometric systems become more widespread, the incentives to misuse or attack biometric systems will grow. Understanding the nature and risk of such attacks will become increasingly important to systems architects, administrators and security managers.

2. Previous models

There are a number of points or vectors where a biometric system can be attacked. While the fake biometric attack has attracted the greatest publicity, other attacks require some form of access to the biometric processing systems and perhaps represent a more significant risk. Some of the early work by Ratha et al. (2001) identified eight possible points of attack (see Fig. 1).

Work by Jain et al., sought to refine this approach. Further work by Wayman (1999) focused on technical testing of

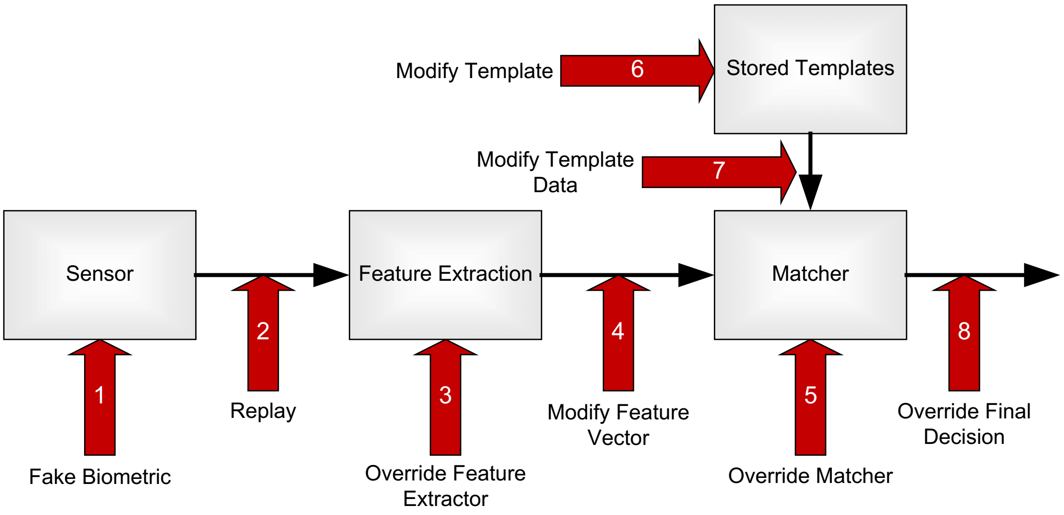


Fig. 1 – Ratha's framework.

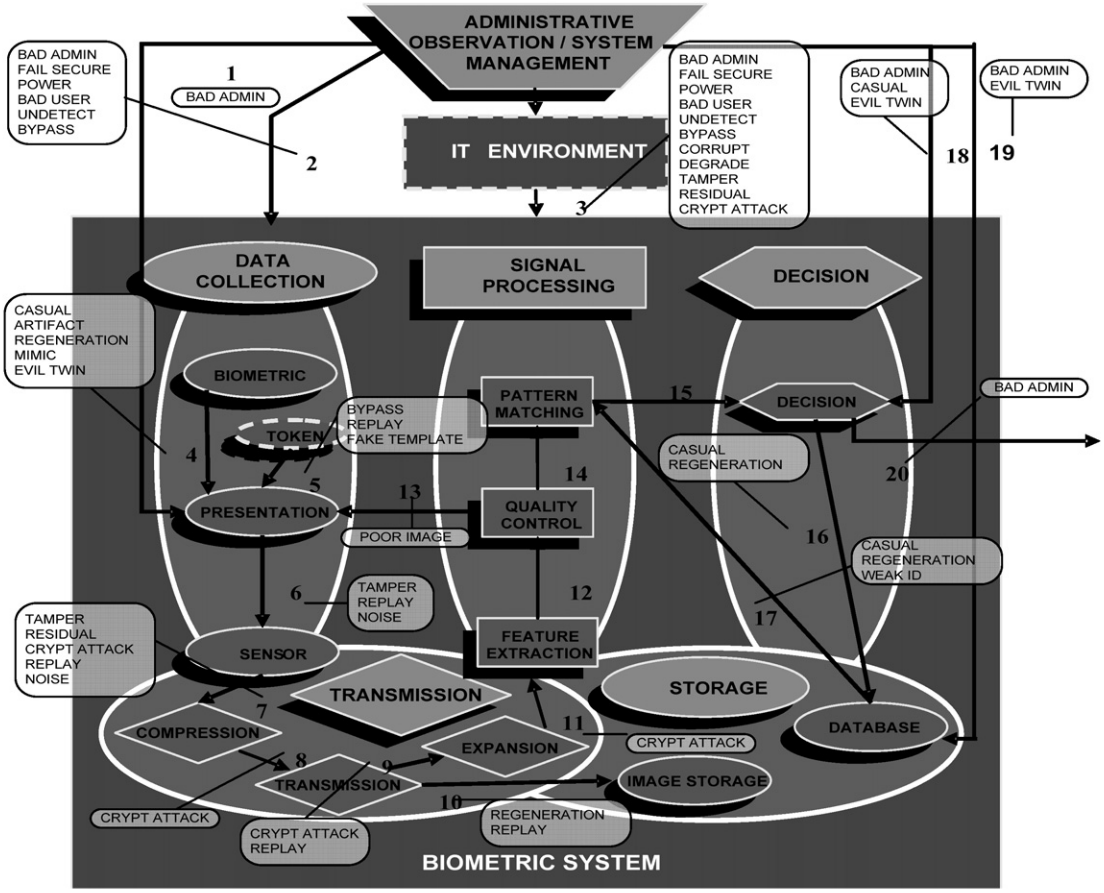


Fig. 2 – Bartlow and Cukic framework.

biometric devices and identified five subsystems, allowing a more refined analysis of potential attack vectors. Bartlow and Cukic (2005a,b) extended this research in a framework combining elements of previous work and adding three components: administrative supervision, IT environment and token presentation. The resultant framework identified 20 potential attack points with 22 vulnerability possibilities (See Fig. 2).

3. A practical view

Attempting to illustrate attack vectors using the frameworks referenced above presents considerable challenges due to the multi-dimensional nature of attacks. These models have become increasingly complex and consequently their utility for practitioners has been reduced.

These models have also not fully accommodated risk-based approaches adopted by many organisations. In order to simplify the analysis of risk of attacks on biometric systems, three dimensions are examined, each of which can be separately analysed for risk. Appropriate risk-reduction and countermeasures can then be selected to manage the risks identified. Finally, the separate risk analyses can be merged to develop a system protection profile.

With adaptation, this approach may also be usefully applied to other technology systems, its utility not being confined to biometric systems.

3.1. *Threat dimensions*

There are three key dimensions of systems' attacks, each of which may require different treatment. These are:

- threat agents;
- threat vectors; and
- system vulnerabilities.

Given the complexity of interactions and the difficulty in illustrating all three dimensions in a single diagram, this paper presents each attack dimension separately. This approach assists in rationalising defences as countermeasures can then be grouped, thus facilitating system management. This approach also facilitates the assessment of risk associated with the threats and threat vectors.

3.2. *Threat agents*

An attack is conducted by a *threat agent*, which is defined as a person who, intentionally or otherwise, seeks to compromise the biometric system. There are three categories of threat agents (Biometric Device Protection Profile, 2001) which are listed below:

- *Impostor*: any person who, intentionally or otherwise, poses as an authorised user. The impostor may be an authorised or an unauthorised user.
- *Attacker*: any person or system attempting to compromise the biometric device or system. Motivation may include unauthorised entry or denial of service.
- *Authorised users*: any person or system authorised to use the biometric system but who may unintentionally compromise the biometric device or system. This category caters for unintentional and human error, such as an administrator error in configuring a system.

Threat agents generally have some degree of technical skill. At the lower end of the risk scale, threat agents may lack specific system knowledge and be poorly funded. A greater threat are those skilled, knowledgeable and well-funded threat agents.

Understanding the types of threat agents can assist in developing effective protection measures. It is regularly demonstrated that authorised users and insiders pose as much, or more of a threat than unauthorised users. For example, the 2005 New Zealand Computer Crime and Security Survey (2005) reported that of the organisations who had experienced incidents, 60% experienced incidents from *outside* the organisation but 70% experienced incidents originating from *inside* the organisation. Other surveys have reported similar observations (CSI\FBI annual surveys, 1996 to 2006). These surveys do not differentiate the type of threat agents.

3.3. *Collusion and coercion*

Associated with threat agents is collusion and coercion where legitimate users are pressured in some way, to provide their biometric and access privileges. This can range from social engineering and promises of payment or some other reward to threats of exposure to some real or imagined offence (blackmail). Often reputations can be irrepairably damaged by allegations, however unfounded, and this is a powerful weapon in coercion.

3.4. *Threat vectors*

Threat vectors are the points at which a system can be attacked and are illustrated in Fig. 3 below. This illustration of threat vectors has been adapted from the Biometric Device Protection Profile (2001) published by UK's CESG and the Study Report on Biometrics in E-Authentication by INCITS (2006). Threat vectors are then individually described.

3.5. *Denial of service*

Denial of Service (DoS) attacks are perhaps the crudest of threat vectors. They range from physical damage or power loss to system attacks designed to corrupt or incapacitate the biometric system. Introducing adverse environmental conditions such as heat, light and dust can degrade the performance of sensors and the quality of data. Other forms of attack, such as introducing electrical or radio frequency contamination can also adversely affect data quality. Specific examples may be the use of portable strobe lights against optical sensors, spillage of liquid on sensors and introducing large static electricity charges.

DoS attacks are generally "noisy" in that they are noticed quickly. In some cases, however, the intent is to have the attack noticed in order to create confusion and alarm and force the activation of alternative or exception handling procedures. Seldom used or exercised alternative or backup procedures will, almost inevitably, present greater opportunity for system compromise and are themselves a threat vector.

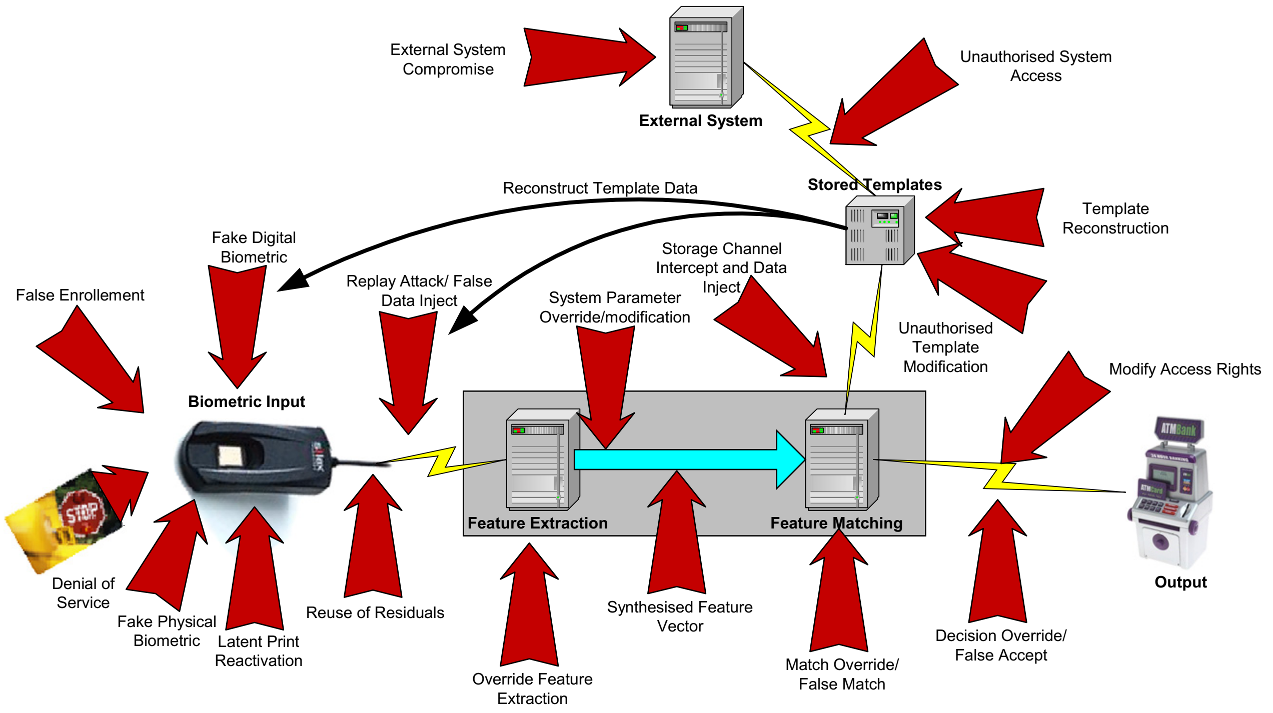

External System Compromise
External System
Unauthorised System Access
Reconstruct Template Data
Stored Templates
Template Reconstruction
Fake Digital Biometric
Storage Channel Intercept and Data Inject
Replay Attack/ False Data Inject
System Parameter Override/modification
False Enrollement
Unauthorised Template Modification
Modify Access Rights
Biometric Input
Feature Extraction
Feature Matching
Output
Denial of Service
Fake Physical Biometric
Latent Print Reactivation
Reuse of Residuals
Synthesised Feature Vector
Override Feature Extraction
Match Override/ False Match
Decision Override/ False Accept

Fig. 3 – Threat vectors.

3.6. False enrollment

The accuracy of the biometric data is founded on legitimate enrollments. If identity is faked, the enrollment data will be an accurate biometric of the individual but identity will be incorrectly matched. This threat vector is seen in other systems, for example, passport applications. Once registered, the system will validate a false identity, and with it any access privileges.

3.7. Fake physical biometric

Perhaps the threat vector that has the greatest prominence when biometric systems are discussed, is spoofing or providing a fake physical biometric designed to circumvent the biometric system. The history of biometric spoofing has been outlined in the introduction to this paper.

This attack can be relatively easily conducted as little or no technical system knowledge is required. The materials for the creation of false biometrics are generally cheap and easily obtainable. Another factor is that these attacks are conducted at the point of entry to the system so many of the digital protection mechanisms, such as encryption and the use of digital signatures, are not effective. Many biometrics (including fingerprints, hand and iris) are subject to this form of attack.

The original biometric can be relatively easily obtained from many sources, with or without the permission and cooperation of the ‘‘owner’’ of that biometric. We leave extensive biometric traces, such as fingerprints and hand prints, on desks, doors, utensils and many other surfaces. Today’s digital camera and digital recording technology has made the acquisition and processing of high-quality images and voice recordings a trivial task.

3.8. Fake digital biometric

A fake digital biometric can have two components outlined below:

- False data using commonly available biometric data such as digital facial images or digitised latent fingerprints. These are sometimes known as masquerade attacks.
- A replay of reference sets. A reference set replay attack takes place *inside* the biometric system and digital defences are more effective here. In addition, the attackers require knowledge of the biometric system and usually also require system access.

3.9. Latent print reactivation

This threat vector is peculiar to fingerprint and palm print scanners. The oils from sweat glands in the skin and residue from touching a variety of surfaces will leave a latent print on the surface of the biometric sensor. These latent prints can be copied or reactivated into readable prints through a range of techniques including powder, the fumes from cyanoacrylate glue, or placing a plastic bag contain warm water over the print.

3.10. Reuse of residuals

Some biometric devices and systems may retain the last few biometrics extracted and templates used, in local memory. If an attacker gains access to this data, they may be able to reuse it to provide a valid biometric. Clearing memory and prohibiting identical samples being used consecutively is an effective defence.

3.11. Replay attacks/false data inject

This category also covers man-in-the-middle attacks. Here the data related to the presentation of a biometric is captured and replayed. Alternatively, a false data stream is injected between the sensor and the processing system. In most cases this will involve some physical tampering with the system. Where templates are stored on an RFID or proximity card, the data are likely to be unencrypted. This can facilitate the unauthorised collection of the data for later replay.

A replay attack is a two or three-stage process, first intercepting or copying the sensor transmission, then possibly modifying the data and finally replaying the signal. Transmission encryption adds a layer of complexity and is an effective defence as the captured signals may be difficult to identify and also must be decrypted, modified and then re-encrypted before replay. Decrypting and re-encrypting data may require the use of specialised tools and the possession of advanced technical skills.

This is also a threat vector for the injection of false data into the biometric system, bypassing the sensor. It is also possible the attacker will automate the intrusion, such as in a ‘‘hill climbing’’ attack (see below).

3.12. Synthesised feature vector

A data stream representing a fake biometric is injected into the system. One approach to generating acceptable data is described as ‘‘hill climbing’’ (Jain et al., 2005; Martinez-Diaz et al.). This technique iteratively changes the false data, retaining only those changes that improve the score until an acceptable match score is generated and the biometric system accepts the false data. This technique requires access to the system’s match scores and communication channels.

3.13. Override feature extraction

This attack interferes with the feature extraction routines to manipulate or provide false data for further processing. Alternatively, this attack can be used to disable a system and create a DoS attack. This is usually conducted through an attack on the software or firmware of the biometric system.

3.14. System parameter override/modification

This threat vector modifies the FAR/FRR or other key system parameters. Adjustments to the system tolerances in feature matching, in particular the false acceptance rate (FAR), can result in system acceptance of poor quality or incorrect data. The US Department of Defense recommends an FAR no greater than 1 in 100,000 and a False Rejection Rate (FRR) no greater than 5 in 100 (Biometrics security technical implementation guide, 2004) for their biometric systems.

3.15. *Match override/false match*

This threat vector could attack software, firmware or system configuration and parameters. Templates are generally unencrypted when undergoing feature comparison and are more susceptible to tampering in this state. The matching decision could be overridden or ignored and replaced with a match. Authorised users are unlikely to notice any anomaly as the system may continue to provide them access.

3.16. *Storage channel intercept and data inject*

Perhaps the attack with the most significant consequences, this attack can compromise both the processing system and any data stored. If the attacker has system access, storage is an easier target as templates are smaller and the data sets less complex than unprocessed biometric data. Examples include the capture of a legitimate template for later use and the injection of a false template. This is an ideal entry point from which to conduct ''hill climbing'' attacks. Successful attacks usually require specific system and template knowledge.

3.17. *Unauthorised template modification*

Templates can be stored on the biometric reader or sensor, on an access card or token or within the biometric system itself. In this threat vector, unauthorised changes are made as templates are modified, replaced or added to the system. Adding an unauthorised template can circumvent any registration procedures and real (but unauthorised) biometrics can be presented and processed by the system alongside legitimate biometrics. A denial of service can be created by corrupting template data or associating users with a modified template. Finally, accidental corruption from a DoS attack, system malfunction or administrative error can also damage template integrity. Loss of template integrity can subvert the identification or authentication processes.

3.18. *Template reconstruction*

One aspect is similar to the synthesised feature vector attack where ''hill climbing'' techniques are used to generate acceptable data. Another form of a template reconstruction attack is scavenge file fragments from data storage. In both these situations, access to the data store is required.

3.19. *Decision override/false accept*

This is a form of bypass attack which ignores any processing and overrides the decision data or injects a false acceptance between the system and the end device (for example a door lock or a cash dispenser). In this case the decision criteria is *accept/accept* in all cases. This may involve some form of physical tampering.

3.20. *Modify access rights*

An unauthorised change to a user's access rights can create a DoS attack when rights are curtailed or alternatively breach security when rights are increased. It is generally achieved by obtaining system administrator rights to enable access to user privileges and other key system parameters and data.

3.21. *System interconnections*

Interconnection with other systems presents at least two more threat vectors, unauthorised (external) system access and external system compromise. If the interconnected system is compromised, it provides an attack vector for the biometric system. Similarly, the communication channel between the systems is open to threat. Often there is little control by the operators of the biometric system over the operation of the external system.

3.22. *System vulnerabilities*

Defects in system design, architecture, production or implementation can all introduce vulnerabilities to biometric systems. In some cases ''secondary'' systems may be integrated into the biometric system and which, if compromised, could leave the biometric system open to exploitation or attack. There are five important areas where vulnerabilities may occur:

- operating systems (server, workstation);
- storage management systems (operating system and application);
- biometric applications;
- sensor software;
- hardware/firmware.

Other key aspects that can be conveniently categorised here include:

- operations management;
- remote management (particularly of FAR/FRR parameters); and
- systems configuration.

These system vulnerabilities are common to many technology systems and have been addressed in some detail in other discussions. It is important to recognise, however, that a system vulnerability can present opportunities for system compromise and the effects can be as equally debilitating as the threat vectors described above.

4. Defences

4.1. *Risk-based approach*

While it is an axiom that defences should be selected for their effectiveness, the criteria for selection are much more difficult to determine. Risk assessment and management frameworks and approaches have been shown to be effective tools in this selection process. The threat dimensions described above are consistent with many of the accepted risk frameworks such as the AS/NZS 4360: *Risk Management* standard, the Treasury Board of Canada Secretariat's (TBS) *Integrated Risk Management Framework* (IRMF, 2001) or the US National Institute of Standards and Technology's *Risk Management Guide for Information Technology Systems*.

The consideration of threats, in relation to risk, provides a threat model which can be used as the basis for architectural designs, information security policy enhancements and security testing plans. Risk analysis is becoming more important as:

- interfaces are standardised;
- specifications and standards become widely available;
- threats to information systems increase;
- consequences of system compromise increase; and
- governance requirements are enhanced.

It is important to recognise that no system can be completely secure and no one single defensive mechanism will comprehensively protect a system. It is also important to recognise that few defensive systems are able to withstand sustained and determined attacks. A risk-based approach to defending systems will allow prudent and pragmatic measures to be identified, can also demonstrate good governance practices and a selection of complementary defences can effectively reduce risk to acceptable proportions.

The vulnerability/robustness ratio of a system can be determined by measuring residual risk, which is generally inversely proportional to the effectiveness of security measures applied.

4.2. *Systems and security architecture*

The two basic architectural decisions in biometric systems are the locations of the biometric matching operations and the template storage. Combined with systems elements, this provides 16 possible architectures (INCITS, 2006). There are also storage alternatives such as Network Attached Storage (NAS), Storage Area Networks (SAN) and other storage arrays. Adding these elements provides 20 possible architectures, each of which should be assessed for risk, threat, vulnerability and then appropriate defensive measures selected (Table 1).

Good practice incorporates proof of concept validation, prototyping and security testing to determine if the architecture and defensive measures selected will provide the required levels of residual risk in the biometric system.

Specific principles incorporated into architectural designs should include the use of "least privilege" and any design should also follow recognised good practice (see *Policy* below).

Table 1 – Architectural combinations

Storage location	Matching location
NAS/SAN/storage array	
Central/distributed (local server)	Server
Local workstation (client)	Local workstation (client)
Device (peripheral)	Device (peripheral)
On-token	On-token

4.3. *Defensive measures*

There are a number of defensive measures that can be taken to minimise the risk of the threat agents, threat vectors and vulnerabilities described above. As with many defensive measures, these are complementary and security should not rely on a single method. Defences can be grouped into six categories and within these groups there are several relevant defensive measures (Liveness detection in biometric systems; Biometrics security technical implementation guide, 2004; Biometric Device Protection Profile, 2001). These are illustrated in Table 2.

4.4. *Challenge/response*

Challenge/response is a technique well-established in protective security. Many will recall or will have used the "Halt! Who goes there?" challenge with a password or pass phrase given in response to the challenge. Today we see this technique applied in many on-line transactions and interactions, such as Internet banking and with utility, credit card and retail organisations. Typically some private reference data are incorporated into the account or transaction set-up and are later used to verify account holders. A classic example is mother's maiden name, although this is well known and an essential piece of information for social engineers seeking to spoof identities.

Challenges can be issued in response to some other "trigger" such as liveness detection failures, lack of movement or changes during the biometric acquisition phase. In biometric systems, users can be challenged, for example, to repeat a particular phrase, blink their eyes, nod heads or present specific fingers to the sensor.

Challenge/response is not restricted to application between the user and the biometric system but can also be used between components of the system. Sometimes described as mutual authentication, it can be an effective defence to replay and data injection attacks, particularly for remote sensors and data storage or other systems' components which are separated geographically.

4.5. *Randomising input biometric data*

A variation of challenge/response is where users are required to enroll multiple biometric samples, such as several fingerprints. Verification will then randomise the sample requested thus adding complexity to any attempt to circumvent the biometric authentication. Such systems may also require multiple biometrics for verification, again adding complexity as any such attempt to circumvent the biometric system will have to prepare several "target" biometrics. This will also assist in defeating attempts to reuse, for example, latent fingerprints on the fingerprint reader.

4.6. *Retention of data*

Generally, sensors are easier to physically access than other biometric system components and are thus more susceptible to attack. In addition, some sensors can store data and copies of templates locally, making them an attractive target.

Table 2 – Defensive measures

	Input device protection	Input data protection	System data protection	Data Storage	System tamper resistance	Secure communications
Challenge/response	✓	✓	✓	✓	✓	✓
Randomising input biometric data		✓	✓		✓	
Retention of data		✓	✓		✓	
Liveness detection		✓	✓		✓	
Use of multiple biometrics		✓	✓		✓	
Use of multi-modal biometrics		✓	✓		✓	
Use of multi-factor authentication		✓	✓		✓	
Use of "soft" biometrics			✓		✓	
Signal and data integrity and identity		✓	✓	✓	✓	✓
Encryption and digital signatures		✓	✓	✓	✓	✓
Template integrity			✓	✓	✓	
Cancellable biometrics			✓	✓	✓	
Hardware integrity	✓	✓	✓	✓	✓	
Network hygiene	✓	✓	✓	✓	✓	✓
Physical security	✓	✓	✓	✓	✓	✓
Activity logging, policy & compliance checking	✓	✓	✓	✓	✓	✓

In most biometric systems, image data are discarded after template generation. Retaining image data may provide a means of resolving spoof claims, although this adds system complexity in dealing with privacy and other storage protection challenges. Clearing data and data buffers are a defence against "man-in-the-middle" attacks and forces an impostor to create data that appear as a biometric sample to the naked eye as well as to the system.

4.7. *Liveness detection*

A key defence to spoofing is "liveness" detection to ensure the biometric sample presented to the reader is from a live person and is not artificial or from a cadaver. Some liveness tests are based on autonomic responses and other can use a challenge/response construct such as blinking an eyelid on command. Liveness detection methods can be incorporated into the biometric reader or can be generated by a separate device. Detection methods include:

- measurement of finger perspiration patterns;
- pulse oximetry where pulse and blood oxygenation are measured by shining a beam of light through the finger tissue;
- skin spectroscopy, which measures the absorption of light by tissue, fat, and blood and melanin pigment;
- photonic and spectrographic measures incorporated into iris recognition;
- thermal measurement;
- head, face, eye and pupil movement;
- synchronising lip movement with voice;
- three-dimensional feature information; and
- printing (dot matrix) and print dye detection.

The use of three-dimensional feature information is considered to improve systems performance against pose and expression variations and changing environmental conditions, such as light and heat Chetty and Wagner. Three-dimensional increases the complexity of the data set by incorporation of subtle variations, particularly in facial images, thus making spoofing extremely difficult. An added advantage is that liveness detection incorporates a non-repudiation element as the user has difficulty in denying that they presented the biometric where there is no evidence of system compromise.

4.8. *Multiple biometrics*

Multiple biometrics increases processing time and adds a level of complexity if more than one biometric is required, for example, a fingerprint and an iris scan. Clearly it is much more difficult to spoof multiple and different biometrics. The requirement for multiple biometrics, however, also adds complexity to the authentication system with requirements such as multiple sensors.

4.9. *Multi-modal biometrics*

Multi-modal techniques are an evolution of multiple biometrics. They can operate using multiple representations of a single biometric or consolidation of multiple features into a new template. Most sensors today will take multiple readings, alternatively, multiple sensors can be used. Processing can range from simple averaging to weighted feature averaging in order to generate match scores. A third technique is to allow biometric sub-systems to individually generate match scores and use majority-voting.

Multi-modal biometrics can assist in improving data quality, precision and integrity, the improved accuracy thus defending against spoofing. It does, however, carry a computational overhead and adds complexity to biometric systems.

4.10. *Multi-factor authentication*

Again similar in concept to randomising data and the use of multiple biometrics, the use of multi-factor authentication, such as a requirement for smart cards, tokens, PINs and passwords, can provide a powerful deterrent to spoofing. It can, however, increase processing time and may reduce the convenience of biometric systems. An attempt to circumvent the biometric system would need both the biometric and the second authentication factor. Multi-factor authentication can be combined with a challenge/response mechanism, further increasing the complexity for any attacker.

4.11. *"Soft" biometrics*

"Soft" biometrics are biometric characteristics which, in themselves, are not sufficiently distinctive to differentiate individuals but in combination provide sufficient data for accurate identification. Examples include age, gender, height, weight, ethnicity and distinctive markings (scars, marks and tattoos). These are the characteristics by which humans identify each other.

This is a defence against spoofing when use in combination with other biometrics. It may also improve systems performance by reducing search times in large biometric databases.

4.12. *Signal and data integrity and identity*

An important component of system integrity is reliable data. Data generated at the sensor must be reliable and it should pass through various stages of comparison and processing with integrity. This is a key defensive mechanism against replay and man-in-the-middle attacks.

Defensive techniques include:

- Time-stamping of the signal between the sensor and the rest of the system. Time stamping, when compared to system clocks or current time, may indicate the use of old or replayed data.
- Use of digital signatures.
- Use of steganography or data hiding (Jain and Uludag). This technique embeds critical data inside another data stream or embeds one biometric data inside another biometric data stream. Such data may include, for example, digital certificates.
- Use of data "watermarks" (Yeung and Pankanti). Again key authentication and verification data can be incorporated into the "watermark".
- Blocking matching attempts where false match thresholds or time periods are exceeded. For example, authorised users are unlikely to have high numbers of false matches in a given time period (with the majority in the morning and at lunch time). Setting limits on the number of attempted matches or number of failed attempts in a given time period, is an effective defence technique.

It is also important that related defensive measures, such as hardware integrity and encryption, are considered.

4.13. *Cryptography and digital signatures*

Encryption of data streams can be an effective defence against data interception and injects. Encryption of data "at rest", such as templates, can be an effective defence against data modification. Digital signatures also defend against data modification for both data in process and "at rest". Key management is an essential component in preserving the integrity of the encryption and digital signature systems. Encryption keys should be secured, preferably not on the biometric system.

4.14. *Template integrity*

The ability to reconstruct biometrics from template data is a concern to privacy advocates and is a threat to template integrity. While many vendors view the template creation process as a one-way algorithm, researchers have shown it is possible to reconstruct sufficient elements from a template to constitute a recognisable biometric. Again "hill-climbing" techniques can be used to iteratively process template data in order to reconstruct a biometric (Bromba, 2003).

A defence against hill-climbing techniques is the use of quantised match scores. This applies rounding techniques to match score calculations in order to minimise differences from small modifications to input images. It thus denies the hill-climbing attack sufficient useful data to identify match score improvements. Soutar (2006) proposes limiting the precision of match scores to make hill-climbing attacks prohibitively time consuming. His research demonstrates unrestricted access to match score data enables a successful attack after a relatively small number of iterations. However, restricting the match score data allows recognition thresholds only after 10^{16} (INCITS, 2006) iterations. This technique limits the effectiveness of a hill-climbing attack.

Some researchers have demonstrated this defence can be defeated but requires extended access to the biometric system in order to be successful, thus increasing the risk of detection. For example, Adler required 122 minutes to process 135,000 biometric comparisons on a PC. While attack techniques and computing power continue to improve, quantised match scores can, at the very least, introduce a significant delay to an attack.

4.15. *Cancellable biometrics*

A characteristic of biometrics is that they are irreplaceable and once compromised, generally cannot be reused. A technique to allow reuse of original biometrics is described as cancellable biometrics (Ratha et al., 2001). This is a deliberate distortion based on a selected transform in which the presented biometric is distorted in the same way at each presentation. The transforms are designed to be non-invertible. Only the transformed data are stored and if these data are compromised, a new transform can be applied, thus replacing the original template.

Cancellable biometrics do not defend biometric systems against attack but will assist in recovery where templates or other biometric data have been compromised. Cancellable biometrics are, however, of little use where the original biometric or image has been compromised.

4.16. *Hardware integrity*

This provides data validation linked to the originating sensor. It may include hardware device identification to generate a unique transaction identification and clearing of local sensor memory to avoid local storage of sensor data or templates.

This can be combined with a challenge/response mechanism or even extended to mutual sensor/server authentication before communication is enabled. Ratha et al. (2001) proposed a pseudo-random challenge to the sensor, the response based on current sensor conditions such as pixel values at selected positions. The response is matched against the biometric data provided by the sensor. This is also a defence against replay attacks.

4.17. *Network hygiene*

As with all technology, good network disciplines and hygiene are essential to the maintenance of system security. Many frameworks and best practice guides are available and apply equally to biometric as well as other technology systems. Examples include ITIL® (IT infrastructure library, 2006), ISO 27005:2005 (ISO/IEC 27001) and COBIT®.

4.18. *Physical security*

Many of the attack vectors described are more easily executed if the attacker has physical access to the biometric system. Physical security, as in many IT security systems, is often the cheapest and most effective deterrent to attempts to circumvent biometric systems. This ranges from physical restrictions to limit access to the biometric readers, to surveillance and guards. Supervised operation or the presence of guards can also defeat other attack types, such as coercion. The risk/reward considerations for attackers should also be factored into the use of physical security as the consequences of discovery and then detention (such as calling the local police), are a significant deterrent to sustained or physical attacks.

Regular inspection and cleaning of equipment is also important. Cleaning, for example, will not only sanitise the equipment for health reasons but also minimises the persistence of latent prints and may improve the performance of the sensor.

Physical security is a key defence in managing access to biometric systems and stored data, such as templates.

Other important physical protections includes items such as:

- tamper switches on sensors and readers;
- alarmed and locked panels for devices and communications interfaces (patch panels etc.);
- protect cabling, in conduit if necessary. Pay particular attention to cabling in non-protected areas, such as ceiling or floor cavities;
- monitored CCTV coverage for readers;
- limited access to readers and sensors, including turnstiles or other forms of physical access control to limit numbers able to access sensors at any one time. This may assist in preventing "tail-gating" or "piggy-back" attacks where the biometric system is used to control access and entry.

4.19. *Activity logging*

Where strong defensive measures are in place, determined attackers may conduct reconnaissance or run the attack over several days or even months, in order to gather sufficient information for an effective systems compromise. Activity logging and pattern extraction can be a useful tool in identifying such reconnaissance or attacks.

In addition to activity logging and monitoring, biometric systems should monitor specific activities and related security events including:

- communication errors from sensors and readers;
- false readings;
- repeated failed authentication attempts.

4.20. *Policy*

Policy is the fundamental framework of security systems. It is a statement of expected behaviours in support of the organisation's objectives. Without a clearly defined security policy, organisations often lack direction, security measures are ineffective and perform below expectations (Cybersecurity operations handbook, 2003) in relation to the security and integrity of their information systems.

Good policy, on the other hand, enhances security and will act as a deterrent to unwelcome, inappropriate and malicious behaviours.

There are several generally accepted standards and frameworks for the management of information security, issued by standards, professional and security organisations. These include:

- ISO 27001, *Information Security Management Systems*;
- BS 7799 Parts 1,2 and 3, *Information Security Management Systems* (Information Security Standard);
- ISO 15408, *Common Criteria* (Evaluation criteria for IT security);
- various NIST Computer Security Publications (Computer Security Resource Center);
- COBIT®;
- IETF (RFC 2196, *Site security handbook*).

4.21. *Compliance checking*

Compliance checking and security assessments play a very important role in:

- maintaining information systems security;
- identifying and facilitating changes necessary to respond to rapidly changing technologies and threats;
- demonstrating prudent governance of information systems; and
- demonstrating compliance with legislation and regulation.

Good compliance systems support risk management systems and decision making. They have close correlation and are complementary to quality control systems. Some compliance tools, such as Nessus (Nessus vulnerability scanner), can monitor technical compliance to assist in keeping systems current and patched against known vulnerabilities and also monitor systems against defined security policies.

5. Conclusion

Much of the activity in spoofing biometric systems has, up until now, been confined to researchers. However, as the use of biometric systems become more widespread, the incentives to misuse biometric systems will also grow. The application of biometric systems in access control and authentication, coupled with uptake by the financial and banking sectors will undoubtedly see an increase in misuse and attacks on biometric systems.

This growth phenomena is not unique to biometrics and has been replicated in many other systems which seek to safeguard information and money.

An holistic approach should be taken when considering any biometric system. It is also important to ensure security is incorporated into the design and architecture from inception. This assists in properly understanding risks and appropriately selecting and implementing defences, in order to avoid those embarrassing and costly security breaches.

The approach presented in this paper accommodates organisational requirements to undertake risk-based analyses and systems security. It is a practical approach to the difficulty of analysing a multi-dimensional threat environment by allowing separate analysis of threat agents, threat vectors and system vulnerability. These separate analysis then draw together system defences, selected for their risk reduction properties, to produce a demonstrably risk-based system protection profile.

REFERENCES

Adler Andy. Reconstruction of source images from quantized biometric match score data. University of Ottawa, <http://www.wvu.edu/~bknc/2004%20Abstracts/Reconstruction%20source%20images%20from%20quantized.pdf> [accessed 25.11.05].

AS/NZS 4360:2004 risk management, Standards New Zealand, <http://www.standards.co.nz> [accessed 01.09.06].

Bartlow Nick, Cukic Bojan. The vulnerabilities of biometric systems – an integrated look and old and new ideas. Technical report, West Virginia University; 2005a.

Bartlow Nick, Cukic Bojan. Biometric system threats and countermeasures: a risk-based approach. In: Biometric Consortium Conference, <http://www.biometrics.org/bc2005/Presentations/Conference/2%20Tuesday%20September%2020/Tue_Ballroom%20B/Cukic_Threats%20and%20countermeasures.pdf>; September 2005b.

Biometric Device Protection Profile, UK Government Biometrics Working Group, Draft issue 0.82-5, <http://www.cesg.gov.uk/site/ast/biometrics/media/bdpp082.pdf>; September 2001 [accessed 13.10.06].

Biometrics security technical implementation guide version 1. Release 2. Defense information systems agency for the US department of defense, <http://csrc.nist.gov/pcig/STIGs/biometrics-stig-v1r2.pdf>; 23 August 2004 [accessed 13.09.05].

Bromba Manferd. On the reconstruction of biometric raw data from template data, M.U.A. Bromba, Bromba GmbH <http://www.bromba.com/>; July 2003 [accessed 14.08.06].

Check Body, Thalheim Lisa, Krissler Jan, Ziegler Peter-Michael. Biometrie (Translated from the original German by Robert W. Smith) c't magazine 2002;114. <http://www.heise.de/ct/english/02/11/114/> [accessed 05.02.06].

Chetty Girija, Wagner Michael. Audio–video biometric systems with liveness checks, University of Canberra, <http://pixel.otago.ac.nz/ipapers/24.pdf> [accessed 03.09.06].

Clarkson University Engineer Outwits High-Tech Fingerprint Fraud, Clarkson University, <www.yubanet.com/artman/publish/printer_28878.shtml>; 10 December 2005 [accessed 19.12.05].

COBIT®, Information Systems Audit and Control Association®, <http://www.isaca.org/> [accessed 10.09.06].

Computer Crime and Security Survey, University of Otago, <http://eprints.otago.ac.nz/342/01/2005NZComputerCrimeAndSecuritySurveyResults.pdf>; 2005 [accessed 08.09.06].

Computer Security Resource Center, National Institute of Standards and Technology, <http://csrc.nist.gov/> [accessed 10.09.06].

CSI/FBI annual surveys, computer security institute, 1996 to 2006, <http://www.gocsi.com>.

Cybersecurity operations handbook. 1st ed. Rittinghouse and Hancock: Elsevier Digital Press, ISBN 1-55558-306-7; 2003.

Evaluation criteria for IT security – Parts 1, 2 & 3: International Organization for Standardization, <http://www.iso.org> [accessed 10.09.06].

Harrison Ann. Hackers claim new fingerprint biometric attack. SecurityFocus, http://www.securityfocus.com/print/news/6717, 13 August 2003 [accessed 13.08.06].

Information Security Management Systems, International Organization for Standardization, <http://www.iso.org> [accessed 10.09.06].

Information security standard. BSI management systems, <http://emea.bsi-global.com/InformationSecurity/Overview/index.xalter> [accessed 10.09.06].

Integrated risk management framework (IRMF), the treasury board of Canada secretariat (TBS), <http://www.tbs-sct.gc.ca/pubs_pol/dcgpubs/RiskManagement/dwnld/rmf-cgr_e.pdf>; April 2001 [accessed 01.09.06].

ISO/IEC 27001:2005, Information technology – security techniques – information security management systems – requirements, <http://www.iso.org> [accessed 10.02.06].

IT infrastructure library, Hompage, <http://www.itil.co.uk/> [accessed 10.02.06].

Jain Anil K, Uludag Umut. IEEE transactions on pattern analysis and machine intelligence, vol. 25, No. 11, November 2003. <http://biometrics.cse.msu.edu/Publications/SecureBiometrics/JainUludag_HidingBiometrics_PAMI03.pdf> [accessed 08.09.06].

Jain Anil K, Ross Arun, Uludag Umut. Biometric template security: challenges and solutions. In: Proceedings of the 13th European signal processing conference (9EU-SIPCO). Turkey: Antalya, <http://biometrics.cse.msu.edu/Publications/SecureBiomet rics/JainRossUludag_TemplateSecurity_EUSIPCO05.pdf>; 2005 [accessed 03.09.06].

Jain Anil K, Pankanti Sharath, Prabhakar Salil, Hong Lin, Ross Arun, Wayman James L. In: Proceedings of international conference on pattern recognition (ICPR) Cambridge, UK, Aug. 2004. Michigan State University/IBM T. J. Watson Research Center/DigitalPersona Inc./Siemens Corporate Research/West Virginia University/San Jose State University. <http://biometrics.cse.msu.edu/icprareareviewtalk.pdf> [accessed 05.02.06].

Liveness detection in biometric systems, Biometrics information resource, <http://www.biometricsinfo.org/whitepaper1.htm> [accessed 05.02.06].

Martinez-Diaz M, Fierrez-Aguilar J, Alonso-Fernandez F, Ortega-Garcia J, Siguenza, JA. Hill-climbing and brute-force attacks on biometric systems: a case study in match-on-card fingerprint verification, Universidad Autonoma de Madrid, <http://fierrez.ii.uam.es/docs/2006_ICCST_HillClimbingAttackMoC_Martinez.pdf> [accessed 03.09.06].

Matsumoto Tsutomu, Matsumoto Hiroyuki, Yamada Koji, Hoshino Satoshi. In: Proceedings of SPIE. Optical security and counterfeit deterrence techniques IV, vol. #4677. Japan: Graduate School of Environment and Information Sciences Yokohama National University, http://cryptome.org/gummy.htm; 24–25 January 2002 [accessed 29.09.05].

Nessus vulnerability scanner. Tenable network security, <http://www.nessus.org/index.php> [accessed 10.09.06].

Ratha NK, Connell JH, Bolle RM. Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal (3), http://domino.research.ibm.com/tchjr/journalindex.nsf/a3807c5b4823c53f85256561006324be/dd12e71773f23bcb85256bfa00685d76?OpenDocument; 2001;40 [accessed 01.09.06].

Risk Management Guide for Information Technology Systems. Special publication 800-30, National Institute of Standards and Technology, <http://csrc.nist.gov/publications/nistpubs/800-30/sp800-30.pdf> [accessed 01.09.06].

Site security handbook, RFC 2196, Internet engineering task force, <http://tools.ietf.org/html/rfc2196> [accessed 10.09.06].

Soutar Colin. Biometric systems security, Bioscrypt Inc., <http://www.silicon-trust.com/pdf/secure_5/46_techno_4.pdf> [accessed 03.09.06].

Study report on biometrics in E-authentication Ver 0.2. InterNational Committee for Information Technology Standards, <http://www.incits.org/tc_home/m1htm/2006docs/m1060112.pdf>; February 2006 [accessed 08.09.06].

Wayman JL. Technical testing and evaluation of biometric devices [Michigan State University]. In: Jain AK, Bolle R, Pankanti S, editors. Biometrics – personal identification in networked society. Kluwer Academic Publisher, <http://www.cse.msu.edu/~;cse891/Sect601/textbook/17.pdf>; 1999.

Wills David, Lees Mike. Six biometric devices point the finger at security. Network Computing, http://www.networkcomputing.com/910/910r1.html 1 June 1998 [accessed 29.01.06].

Yeung Minerva M, Pankanti Sharath. Verification watermarks on fingerprint recognition and retrieval, <http://www.research.ibm.com/ecvg/pubs/sharat-water.pdf> [accessed 08.09.06].

Chris Roberts is a qualified Chartered Secretary, Management Accountant and Information Systems Auditor. He has over 35 years of IT, commercial and audit experience and has specialised in information security and assurance over the last 18 years. Chris has also worked extensively in the areas of e-fraud, other IT related investigations and computer forensics. He has provided specialised assistance to a wide variety of government, state-owned enterprises, financial and other private sector organisations. He has also worked extensively with international AID organisations such as USAID, CIDA, UNDP, SIDA and The World Bank on sponsored projects for government ministries, departments and state-owned enterprises in several countries.

ELSEVIER

Computers & Security

Information Lifecycle Security Risk Assessment: A tool for closing security gaps

Ray Bernard

Ray Bernard Consulting Services, USA

Keywords:
Data lifecycle risk analysis
Electronic data security
Electronic document management
Enterprise data management
Information lifecycle security risk assessment
Information security risk assessment
Physical data security
Proprietary information protection
Records and information management

ABSTRACT

News media continue to report stories of critical information loss through physical means. Most information security programs include physical protection for information system infrastructure, but not for the physical (non-electronic) forms of the information itself. Thus organizations have persistent critical information vulnerabilities that are not addressed by even the most extensive of information systems security programs.

An *Information Lifecycle Security Risk Assessment*, as described in this paper, can be used to extend the reach of information security programs to encircle all forms of critical data from creation to destruction—even data in human memory form. Such an assessment can leverage existing data management and information systems security efforts. By incorporating both electronic and physical information elements, previously unaddressed information security gaps can be identified and mitigated. The end result should be a risk treatment plan which senior management can understand and approve, and which managers and security personnel can execute.

A high-tech manufacturing company pointed to a $10 million drop in service business revenue as evidence of substantial quality improvements in their product lines. An astute board member launched her own investigation and determined that the real cause was encroachment on the service business by competitors, who had been illegally obtaining physical copies of proprietary company information and for over two years, had been using it to quietly take over customer service accounts.

For over a year the largest sales branch of a national company experienced a level of sales competition unheard of in any other sales office, resulting in the lowest sales closing average in the company's history. Personnel from a competing company were sneaking up to the sales team's conference room window at night, and peering through tiny slots in the window blinds to copy the daily list of hottest sales prospects from the white board—including products and anticipated sales amounts.

While incidents of information loss by physical means of one kind or another are routinely reported by news media and security publications, many instances—like the two described above—are not publicly disclosed. For decades *Records and Information Management* (RIM) professionals have managed information in paper or other physical forms, and have utilized physical security programs to manage the protection of that information. Then how is it that today many critical information losses are the result of successful *physical* attacks? How is it that the protection of information in physical forms is poorly addressed in many organizations, despite the increased awareness of the importance of information protection?

1. Information security redefined

Today broadband networks and high-capacity electronic data storage technologies enable organizations and individuals

E-mail address: raybernard@go-rbcs.com

to create, receive, store, access and publish information in quantities—and at speeds and economies—that remain impossible with physical forms of data. Organizations have embraced electronic forms of information for their ability to accelerate the pace of any information-based activity. Electronic forms of data have substantially replaced physical forms of data for most organizations.

Thus in recent years several new phrases have replaced *Records and Information Management* in organizational parlance: *Electronic Document Management*, *Enterprise Data Management, Enterprise Content Management, Document Lifecycle Management* and most recently *Information Lifecycle Management*. Such approaches are concerned with the practice of applying policies to the effective management of information in all aspects of its useful life. These new approaches to data management reflect the migration from a physical to an electronic organizational data landscape.

However, unlike their predecessor, RIM, most of the solutions under these names have a common focus mainly or solely on the electronic aspects of data handling and storage. For most organizations critical data still exists in other forms, and their security is not addressed by the security components of the electronic data management approaches.

Information systems security practitioners are aware of the fact that the scope of their work is limited to electronic data. For example, the CISSP designation stands for Certified Information Systems Security Professional, where ''Information Systems'' means ''electronic data systems''.

In contrast, the well-known information security standard, ISO/IEC 17799:2005, states in its introduction:

> Information can exist in many forms. It can be printed or written on paper, stored electronically, transmitted by post or using electronic means, shown on films, or spoken in conversation. Whatever form the information takes, or means by which it is shared or stored, it should always be appropriately protected.
>
> Information security is the protection of information from a wide range of threats in order to ensure business continuity, minimize business risk, and maximize return on investments and business opportunities.

Yet in everyday use the term *information security* is most often applied to *electronic* information security, the realm of IT security practitioners, where the application of physical security is limited to information systems physical infrastructure. This amounts to an unintentional redefinition of *information security*, causing vulnerabilities to many non-electronic forms of data to fall out of organizational view.

Another part of the picture is the fact that regardless of the inclusion of physical and environmental security in 17799 or any other information security standard, the vast majority of information security practitioners have neither the knowledge nor the means to implement physical security controls for non-electronic forms of data. For information security to be complete, all forms of data must be addressed, and they must be addressed by the personnel who have the knowledge and means to identify and mitigate their information security risks.

2. Infrastructure focus also afflicts physical security

A close focus on infrastructure also can be found with physical security practitioners. They attend to building external and internal structures, means of access (such as doors, windows, roof hatches, etc.) and physical facility vulnerabilities. Physical forms of information are protected in part as a side-effect of protecting the rooms that contain them, similar to how electronic information is protected in part by physical protection of information systems infrastructure. Outside of government and private sector facilities with classified information, in most companies many physical forms of information are not subject to sufficient security controls. The exceptions are generally those organizations that have suffered a physical loss of critical information, and have closed the open doors related to their loss events. Information is usually only loosely tied to physical protection zones, and that is done at the time that physical protective measures are initially established. As organizations change, their usage of information changes, and the physical forms of information and their locations change also. Yet physical protective measures are rarely reevaluated unless physical building structures change.

3. Need for a workable process

What is lacking is single a process whereby the critical information assets, *in all of their forms*, can be identified, cataloged, ranked in terms of their criticality,[1] and protected by establishing and maintaining suitable controls. The solution involves what is often referred to as *security convergence*: the collaboration between IT Security departments and Physical Security departments, groups which are historically separate functions in most organizations. To date what has helped to keep the two groups separate is their infrastructure focus, which takes them in different directions. What does enable the two groups to collaborate successfully is the *risk perspective*. It provides a common vision that makes a single process workable for both groups, and can encompass both physical and electronic forms of information. The information lifecycle approach provides a birth-to-grave scope that facilitates identifying all of the forms that information can take, electronic and physical. This results in an information risk assessment process that is truly complete in its scope.

4. Information security stakeholders

Collaboration between Physical Security and IT Security departments is only a starting point for an Information Lifecycle Security Risk Assessment. To be successful the risk assessment process must involve personnel outside of the security departments. Typical information security stakeholders include personnel from Human Resources, Legal,

[1] Criticality is the severity of impact of the loss of the asset.

Compliance, Audit, and Risk Management; but they can only provide part of the risk picture.

Users of the information in various business units understand its role in their critical functions, and the impact of its loss. They also know how information is accessed (not always in conformance with policy), what forms the information can take, and where physically the various forms of information can be located.

Managers who are responsible for the business units that depend on the information assets are information security stakeholders from several perspectives. First, they often make decisions about who can access information, and where. Second, they have a responsibility to see that the information assets on which they depend are safeguarded, and so require input into the security process at least in terms of identifying the critical assets. Third, they must also support and enforce security policy and controls within their own areas, which is an organizational responsibility. Fourth, sometimes security considerations warrant a change to a business processes—at times requiring the physical relocation of a business function that is not in a secure enough location. Such changes require not only management approval but also active management involvement to execute them successfully.

Additionally, senior management must be informed about and provide approval of major security initiatives. After all, the information assets are corporate assets—not the security departments' assets—and the decisions about what levels of risk to accept are not security department decisions. Security departments and security executives can and should make recommendations, but ultimately the decisions must rest with the senior executives who are responsible for the corporate assets, and with the executives who are responsible for success of the business units that depend on the assets. Thus senior executives are also information security stakeholders.

Generally senior managers usually do not understand all of the workings of security, but they do not need to. When presented with a good risk picture and prioritized risk treatment plan, they can easily weigh the cost of risk reduction measures against the potential impact of risk events on the business. This information allows them to become security advocates for the assets they are responsible for or on which they depend. Strictly speaking this is not an advocacy on behalf of security; it is an advocacy on behalf of the business.

5. Collaboration strategy

The strategy proven to be most successful for fully addressing the critical information risk picture is one that involves all of the stakeholders: an *Information Security Risk Management Council*, a group whose actual name will vary (task force, committee, etc.) depending upon the organization. Such a council can fulfill its role when its members can speak to the full lifecycles of the critical information assets, either by their own direct involvement or by survey and discussion with those who are directly involved. Early in its work a council may discover that its members cannot adequately address the full lifecycle of all of the critical information assets. It is usually a simple matter to expand the membership slightly to achieve that coverage. Typically the council includes members from Human Resources, Legal, Compliance, Audit, Risk Management, IT Security, Physical Security, Corporate Security, and the organization's various business units. The council may report to the CIO, the CEO, the CFO, or to whichever senior executive volunteers for or is assigned top-level oversight. There are usually dotted line reports as well.

6. Information lifecycle

The roles or functions involved in information handling constitute key aspects of the information lifecycle from an analysis perspective. They form a simple checklist that can help guide the effort to identify the various forms information can take:

- Creation and Receipt
- Storage
- Distribution and Transmittal
- Access and Use
- Maintenance
- Disposition and Destruction

Creation and Receipt deal with records from their point of internal origination or their entry into the organization. Information forms can be written, printed, electronic or verbal and include correspondence, contracts, applications, reports, drawings, production or transaction records, and many other forms of data.

Storage refers to all of the places where any form of information is stored, including human memory.

Distribution and Transmittal are processes involved in getting the information to locations where it can be accessed and used. This may happen automatically according to some process or policy, or on request or demand.

Access and Use take place after information is distributed, and may involve converting the data from one form to another, such as printing reports or documents for review, and information sharing on an individual or group basis.

Maintenance is the management of information. This can include processes such as information filing, archiving, retrieval and transfers, as well as changing the classification of information as its value, relevance or validity changes.

Disposition and Destruction involve handling information that is rarely accessed or is required to be retained in specific formats for specific time periods, and is then destroyed by appropriately secure means when it is no longer valuable or required to be retained.

In addition to helping identify the various forms that information can take, there is another beneficial aspect of the information lifecycle approach that pertains to security cost and efficiency. The value of some information changes over time and a lifecycle analysis can identify those change factors. It is good business practice, as well as good security practice, to adjust the level of resources used to safeguard information as the criticality of the information changes.

7. Information Lifecycle Security Risk Assessment

The first step of an Information Lifecycle Security Risk Assessment is to determine or identify:

- the full lifecycle of each operationally critical data asset (creation or receipt, storage, distribution and transmittal, access and use, maintenance, disposition and destruction);
- all the forms in which the data can exist at each point during its lifecycle;
- all the physical locations at which each form can be found or produced;
- what corporate security policies and procedures exist (if any) regarding the various forms of data in each location;
- what personnel (internal and external) can possibly access the data, regardless of whether or not such access would violate any policies that may exist; and
- the effectiveness of any security measures being applied, including inspections and audits.

This provides a baseline picture that can be used to perform a risk analysis and develop a prioritized list of cost-effective measures that should be applied to each data asset during its lifecycle. The remaining risk analysis steps can follow whatever qualitative or quantitative risk analysis methodology is most applicable for the organization. Risk analysis recommendations should include the categories of items shown in Table 1.

An important final step is to update the ongoing information systems risk management program to include periodic checks for changes to each data asset's lifecycle. Change management should trigger reassessment whenever a critical data asset's lifecycle changes.

8. Protecting physical forms of data

Ironically it is the migration away from physical forms of data to electronic forms that makes securing physical forms of information much easier today than it has been in the past. These are some of the reasons:

- Regulations (like Sarbanes–Oxley and HIPAA) require deployment of physical security measures; this is a new driver for physical security.
- Publicized instances of physical loss of critical information have educated senior management to the real dangers of physical security gaps.
- The information security programs and business continuity plans of many organizations have cataloged the critical information assets and provide a significant head-start in the identification of electronic and physical forms of critical information.
- The information classification schemes of enterprise data management programs can be used for the physical forms of information as well, significantly reducing the preparation effort involved in cataloging physical instances of information.
- The framework of an information security management system, such as what ISO/IEC 27001:2005 defines, can also be utilized to establish and maintain physical information security as an incremental effort to existing information systems security management.
- Role Based Access Control implemented for information systems access can be extended to physical access control systems (PACS), either through manual processes and procedures or via integration with an Identity Management System or corporate directory. This provides a way to include physical forms of data in policy-based management of information access, which can be a boon to compliance management

9. Human protective measures

There are some forms of data that can only be protected by appealing to their human custodians. Where information exists in human memory form, security measures like non-disclosure agreements, internal disclosure policies, and

Table 1 – Recommendation categories

Recommendation	Example or explanatory note
Security strategies	Example strategy: for each form that a data asset can take in its lifecycle and for each location where the data form can exist, ensure that a specific person or role is assigned responsibility for the data asset's protection.
Security policies	Policies determine what protective actions are taken when, where and by whom.
Security procedures	Specific and standard steps that implement the actions required by security policies.
Compliance monitoring	Implement compliance monitoring as appropriate by IT security, physical security, or audit department depending upon the security measures to be monitored.
Corporate safeguards	Where significant corporate liabilities exist, institute measures that help safeguard the corporation, for example: forensics quality recorded video surveillance or adjusted insurance coverage.

security awareness training apply. Individual personnel briefings should be utilized when personnel are terminated or transferred, to provide a reminder about information protection obligations that continue despite leaving the organization or some part of it. Some forms of information—like data on PDAs, cell phones and notebook computers, as well as printed information—require safeguarding by the person who has their custody.

Smart card based security systems, especially those which incorporate biometric authentication for critical data, can provide a secure bridge from system to human custody, by controlling the transfer of electronic data into physical form. Printing solutions (such as FollowMe® Printing, Follow & Print, and Cardn'n'Print) exist that require the authorized individual to present a security card and/or a fingerprint at the destination printer before controlled documents will actually print. By applying policies about where sensitive data are allowed and not allowed, printer selection can be restricted by printer location based upon the classification of the data being printed, as well as by the security privilege of the individual printing the information.

Similar restrictions can be applied to writing data to a disc or memory stick, to provide auditable physical chain-of-custody control as information is transferred to a portable medium.

10. Lifecycle approach advantages

One advantage of the lifecycle approach is that it can be applied to any security risk assessment methodology. Its function is to provide a process for the identification of the all of the various forms of information that require protection, which fits into the asset identification step of any information risk assessment methodology. Another advantage that is unique to this approach is that it constitutes a simple point of collaboration in which all security stakeholders can participate, thus providing a bridge between corporate, physical and IT security participants regarding information protection.

Ray Bernard is a security consultant and a writer, who has provided pivotal direction and education in the security profession and the security and building controls industries for more than 20 years. He is the most highly published writer on the subject of the convergence of physical convergence and IT, with more than two dozen full-length feature articles on the subject in a number of security trade publications, in addition to a monthly column, ''Convergence Q&A'', in *Security Technology & Design* magazine. He is a frequent presenter at security conferences and workshops, and conducts a full-day security convergence track each year at the *CardTech-SecurTech* conference. His security consulting clients include Fortune 500 companies, international airports, and critical government and public facilities. Additionally he is the founder of ''The Security Minute'' electronic newsletter, the first newsletter for all security stakeholders. Bernard is also certified as a *Physical Security Professional* (PSP) by ASIS International, and holds the Certified in Homeland Security *CHS-III* designation from the American College of Forensic Examiners Institute (ACFEI).

Computers & Security

Decoding digital rights management

Danny Bradbury[1]

Keywords:
DRM
Piracy
Copy protection
ACSS
Entertainment

ABSTRACT

Digital rights management technology is designed to prevent piracy and facilitate the creation of innovative business models around digital content. Its technological limitations may be surpassed only by its economic ones.

It's not often that you get sued for breaking your customers' computers by installing spyware on them, but that's just what happened to Sony BMG in late 2005.

The company began including software on music CDs that automatically ran when the disk was played on a PC. The software allowed users to copy the music to a hard drive, but only permitted them to write the tunes to three new CDs.

Windows expert Mark Russinovich discovered that the program used spyware-like techniques. It concealed itself from users, damaging Windows if they tried to manually uninstall it. It consumed up processor power by constantly scanning files, and it 'phoned home' to Sony's servers every time a CD is played.

Sony backed down after being sued by consumers in the state of California, and by the Government of Texas. On November 21st, the digital rights advocacy group the Electronic Frontier Foundation stepped up with yet another lawsuit.

This is one of the most visible cases in which content distributors have attempted to use digital rights management (DRM) and seen it backfire. DRM attaches usage rules to digital content to govern its consumption by the user. Advocates argue that it enables companies to stop piracy while wrapping new and innovative business models around their content. A music company may offer a downloadable track that could be played just twice, for example, in a 'try before you buy' model that would require a user to unlock the file by purchasing a key.

DRM advocates should be used to public embarrassment. In 1998, the content distribution industry came together to create the Secure Digital Music Initiative (SDMI). SDMI was a technology designed to wrap rights protection around digital music content. Two years later, in a move designed to demonstrate the invulnerability of the system, the SDMI forum issued a challenge to hackers, inviting them to try and break the system.

Inevitably, someone did; Princeton professor Ed Felten led a team that exposed basic flaws in SDMI's design. Felten announced his intention to present a paper detailing his methods, and the SDMI forum reportedly threatened to sue him under the Digital Millennium Copyright Back (DMCA), which had been signed into law in 1998. The Association of Computing Machinery in the US testified in Felten's defence.

The DMCA is something of a recursive law, designed to protect the very act of protection. It implements the copyright treaty outlined by the World Intellectual Property Organisation, and two of its provisions are particularly relevant to DRM. Firstly, it forbids the circumvention of technological measures used by copyright owners to protect their works. Secondly, it prohibits tampering with copyright management information.

E-mail address: danny@itjournalist.com

[1] A freelance journalist.

"You might get a DRM that lets you play some content twice," explains David Fewer, staff counsel at the Canadian Internet Policy and Public Interest Clinic (CIPPIC), an organisation established at the University of Ottawa in 2003 to ensure balance in the policies surrounding new technologies.

If you wanted to play the content more than twice, you might decide to manipulate the underlying technology, he explains. "That would be a violation of the anti-tampering rules, and if to do that you had to hack through some security, that would violate the anti-circumvention measures".

Industry commentators such as Peter Eckersley, staff counsel for the anti-DRM group the Electronic Frontier Foundation (EFF), say that the DMCA was written at a time when the content industry believed that its existence was threatened by digital piracy. Indeed, the protection of the content was the original motivation for creating DRM in the first place, argues Princeton's Felten.

The idea was that companies would be able to capitalise on the digital distribution of content (with the inherent logistical savings facilitated by easier distribution), without sacrificing revenue to authorised copying. "It's controlling the content owner's fear that things will go the way of video in China," confirms Corey Ferengu, VP of solutions marketing and strategy at Macrovision.

However, companies attempting to develop and implement DRM seemed to face an eternal problem: thus far, many DRM mechanisms have been successfully broken. Aside from SDMI, the other landmark case was the Content Scrambling System (CSS), the copy protection system used for DVDs, which also dictated the right to play DVDs in certain geographical regions. The technology was cracked in 1999, leading to the development of tens of different open source DVD players and software 'rippers'.

It is technically possible to produce an unbreakable DRM, argues Ferengu, but it would be commercially useless. "It would be so hard for people to use it. We could put rip protection in a DVD today but the number of players that could play it would come down drastically," he says. "What you're really trying to do is prevent the casual user from copying."

But this motivation for using DRM is weakening all the time, claims Felten. Given that most if not all DRM must be breakable in order to be commercially viable, it seems likely that someone will always break commercial DRM. And the ubiquity of file sharing networks means that once one person has broken the DRM on a file, it will spread quickly. Moreover, music is routinely sold on CDs, few of which are copy-protected. Ripping this music and sharing it online makes the whole point moot.

Consequently, Felten says that the motivation for using DRM is changing. "Companies are realising that they can shape the market rather than stop infringement," he says. Even if everyone comes to admit that it will not stop infringement, we'll see DRM used for this reason. We're in that transition now."

He worries that companies will use DRM to gain an advantage over the competition. If a company can use DRM to make all of its own products work together while excluding the products of its competitors, then it can increase its market share by building an ecosystem tying together distribution and the consumption of content.

Apple, which uses a proprietary DRM system called Fairplay, is a clear example of a company that uses DRM to maintain market control, suggest commentators. "Any DRM you design will have someone in charge of deciding who is in charge of operation and who isn't," says Felten. "Once you do that, you know that whoever has that control will use it for their own business purposes. Apple is in control of Fairplay in this sense."

Those who consider such opinions politically biased might consider Apple's victory over the French legislative system earlier this year, in which it prevented the country from mandating the sharing of rights management technologies with rivals in a bid to promote interoperability. At that time, the company called the proposed law "state-sponsored piracy".

Putting control of the way that content is used in the hands of the people that distribute it is inherently undemocratic, argues Fewer. "It's a way for companies to write their own law," he complains. "You write a use policy, and copyright is irrelevant to the use policy. It's about what the distributor wants."

In his book Free Culture, Stanford law Professor Lawrence Lessig uses the example of the Adobe e-book reader, which he says disallows any printing or copying of the text of Aristotle's *Politics*, for example, even though the text is in the public domain and not subject to copyright law at all. "This is the future of copyright law: not so much copyright law as copyright code," he says.

There are nevertheless attempts to solve the interoperability problem. The Open Mobile Alliance (OMA) an industry consortium of companies related to the mobile phone industry, has approved version 2 of the OMA DRM standard. A PKI-based system, OMA DRM 2.0 logically separates content from the usage rights information that facilitates its consumption. They can be obtained separately, possibly from different providers. The rights management information is contained in a Rights Object, delivered as an XML document, which is then cryptographically tied to one or more devices. Associating the rights object to multiple devices is an important part of the value proposition, because it enables content to be shared between devices (say, a mobile phone and a PC), as long as the distributor allows it.

"From a technical point of view it is very rigorous and well throught through," says Al Hawtin, VP of sales and marketing at Elliptic Semiconductor, which specialises in creating electronic components that can enforce DRM. "In the next year we could see some major cellphone network vendors rolling out DRM schemes based on OMA and that may set the foundation for the beginning of the standardisation process. It's only that that will break Apple's bubble," he says.

While the OMA works to enforce interoperable DRM in the mobile sector, vendors of home entertainment media are working on their own schemes. HD-DVD and Blu-Ray are two competing standards for high-capacity optical disk storage, and are essentially the successors to the conventional DVD. Determined not to make the same mistake that was made with CSS, the groups behind these standards are trying to enforce more rigorous copy protection to underpin DRM.

Both formats will use the Advanced Access Content System (AACS), which differs from the CSS protection mechanism used in traditional DVDs in that it is a dynamic protection system. Under CSS, DVD players created by a licensed manufacturer were shipped with a set of cryptographic keys. These keys were designed to decrypt content on any future DVD media that the user chose to play. The idea was that, if a DVD player was found to have been compromised, future DVDs could be mastered with content that was incompatible with that players keys, meaning that it could not play any new content.

CSS had two main flaws. Firstly, because the same keys were shipped across broad numbers of players from individual manufacturers, it proved impossible to disable a set of keys, because that would render an entire range of DVD players unusable. Consumers wouldn't have stood for this. The second flaw was that the proprietary cryptographic keys used for the system were relatively weak, making it easy to compromise a player.

The AACS group hopes to learn from the mistakes of the Copy Protection Technical Working Group (CPTWG) that produced CSS. It uses 128-bit AES encryption for its keys, making them much tougher to crack. It also uses a more granular key distribution system, giving each player its own unique set of keys which are managed using a binary tree. Media are shipped with a Media Key Block, which compliant devices can process with their own device keys to produce a Media Key.

This process enables a compromised device to be 'lopped off' the tree by shipping all future media with a new Media Key Block designed not to work with a compromised key. It will allow all devices to compute the same Media Key other than the compromised device, which will compute a different key and fail to play the content.

The AACS system relies on the storage of keys on the physical medium that cannot be read by a consumer player, thus preventing bit-by-bit copying. But that copy protection, on which DRM can then be built, relies on the physical medium for its execution. As content becomes increasingly divorced from the physical means of distribution, companies are inventing other ways to include DRM in the mix.

Adobe's LiveCycle Policy Server, for example, requires its reader to 'phone home' across the Internet so that the server can authorise the user to consume the content displayed by the reader according to the rules set by the content distributor. The distributor can also grant the reader a certain number of offline accesses, which according to Adobe's senior sales engineer David Stevenson enables access when the reader is offline.

The other development that promises to maintain DRM while divorcing the content from the physical media is the Trusted Platform Module, an integrated circuit designed to be tamper-proof, which can store information securely.

"Your computer is no longer quite under your control. It has this bit of hardware built in that isn't loyal to you, it's loyal to someone else," says the EFF's Eckersley. One feature of the TPM is remote attestation, which allows third parties to check that the configuration of the computer is the same, and that nothing has been tampered with. This could be used to ensure the integrity of a player with a DRM wrapper that enforced a company's distribution rules, for example. However, the use of TPMs for DRM purposes has been very limited to date.

Meanwhile, it seems likely that provisions for a broadcast flag will appear in a version of the Telecommunications Act 2006, which US congress had put off until after the November 2006 elections. A broadcast flag would be a signal included in a digital TV broadcast that would tell personal video recorders (PVRs) what content was copy protected and what wasn't. A mandate requiring all digital TV devices to be broadcast flag-compliant would lock down the redistribution of televisual content between different devices, effectively coding rights management into the signal and the player.

As tensions between content distributors and consumer groups continue to grow, the quest for an invulnerable DRM technology continues. As the battle drags on, it is becoming clear that DRM is not a technological issue alone: it sits along a complex continuum involving legal, political and economic considerations. No wonder, then, that this is proving such a difficult hurdle for distributors and consumers alike.

Danny Bradbury has been a technology journalist since 1989, covering a range of IT-related issues including security, networking, and software development.

ELSEVIER

Computers & Security

From the Editor-in-Chief

IFIP workshop – Information security culture

Numerous surveys continue to suggest that peoples' attitudes and lack of awareness of security issues are amongst the most significant contributors to security incidents. As such, it is evermore apparent that security can only be effective if staff know, understand, and accept the necessary precautions. This highlights the need to foster a culture in which users are aware of the security issues that pertain to them, and have the required knowledge and skills to act appropriately. With this in mind, the papers in this issue are drawn from a themed workshop on 'Security Culture' that was held as part of the IFIP SEC 2006 conference in Karlstad, Sweden, during May 2006. The workshop was jointly organised by IFIP TC11 working groups 11.1 (Information Security Management) and 11.8 (Security Education) – reflecting that an effective security culture represents one of the necessary foundations for information security management, and cannot be achieved without appropriate attention to security awareness, training and education for staff.

The workshop solicited short paper submissions describing original ideas, research or practical experiences related to the subject of security culture. A call for papers was issued in parallel with that of the main SEC conference, and ultimately received 14 submissions, from which six were accepted for presentation at the workshop, following review by three members of a specially assembled programme committee. The workshop session was well attended, and provoked a range of interesting discussions amongst the participants. Given the quality of the papers and the level of interest it was decided that authors from a subset of the presented works should be invited to further develop their ideas for publication in this special issue of Computers & Security.

Each of the papers presented in this issue has been considerably extended when compared to the corresponding workshop submission (with the original papers having been limited to six pages each). This has given the authors the opportunity to provide greater detail, as well as, in some cases, to update the content to reflect feedback received during the workshop and any recent developments in their research. In addition, to further ensure the quality of the works, each paper has been subjected to further review by TC11-affiliated members of the international board of referees.

Thanks are due to all of the authors that submitted their work, as well as to Prof. Natalia Miloslavskaya and Lieutenant Colonel Ronald Dodge from WG11.8 who assisted with the staging of the original workshop.

Prof. Steven Furnell
Chair, IFIP WG11.1 – Information Security Management
Network Research Group, School of Computing
Communications & Electronics
University of Plymouth, Plymouth, UK
E-mail address: steven.furnell@plymouth.ac.uk

0167-4048/$ – see front matter

doi:10.1016/j.cose.2006.10.012

ELSEVIER

Computers
&
Security

Value-focused assessment of ICT security awareness in an academic environment

L. Drevin[*], *H.A. Kruger, T. Steyn*

North-West University, Private Bag X6001, Potchefstroom 2520, South Africa

Keywords:
ICT security awareness
Value-focused assessment
Fundamental objectives
Means objectives
Security culture
Security planning

ABSTRACT

Security awareness is important to reduce human error, theft, fraud, and misuse of computer assets. A strong ICT security culture cannot develop and grow in a company without awareness programmes. This paper focuses on ICT security awareness and how to identify key areas of concern to address in ICT security awareness programmes by making use of the value-focused approach. The result of this approach is a network of objectives where the fundamental objectives are the key areas of concern that can be used in decision making in security planning. The fundamental objectives were found to be in line with the acknowledged goals of ICT security, e.g. confidentiality, integrity and availability. Other objectives that emerged were more on the social and management side, e.g. responsibility for actions and effective use of resources.

1. Introduction

There is a wide range of threats to information security, e.g. human errors, acts of sabotage, theft, forces of nature, technical failures, etc. According to Whitman and Mattord (2005), employee errors are rated among the top threats to information assets, and security education, training and awareness form part of the process to educate staff on information security. Pfleeger and Pfleeger (2003) state that people using security controls must be convinced of the need for it. They have to understand why security is important in a given situation. Czernowalow (2005) reports on the views of the managing director of a big IT company who looks at security from a business point of view and states that a single case of abuse can cause more costs than the establishment of a security system. The perception is that the cost of training employees is less than the potential penalties incurred if legislation was not adhered to or the company's systems were attacked. Employees should know the rules otherwise they cannot be expected to follow them. Training can prevent staff from accidentally acting inappropriately. Effective use of security controls that are in place can only be achieved when employees are aware of the need for security. BS7799:1 has a section on user training and the objective is to 'ensure that all users are aware of information security threats and concerns, and are equipped to support organizational security policy in the course of their normal work' (BS7799, 1999).

It is necessary to do an assessment to measure the awareness of staff members regarding information communication and technology (ICT) security in general. Focus areas are necessary to measure relevant areas otherwise many aspects can be looked into without getting to the real shortcomings or issues. The value-focused thinking method, explained in Section 2, was used in a university environment as part of a bigger security awareness project.

The aim of this paper is to introduce the approach of value-focused thinking as applied to ICT security awareness. This paper will first discuss the value-focused thinking method of arriving at fundamental objectives to identify important security awareness aspects. Next, the security awareness project

[*] *Corresponding author*. Tel.: +27 18 299 2531; fax: +27 18 299 2557.
E-mail addresses: ldrevin@acm.org (L. Drevin), rkwhak@puk.ac.za (H.A. Kruger), rkwts@puk.ac.za (T. Steyn).

will be discussed after which the results obtained will be given. A discussion of the fundamental objectives will follow and lastly a conclusion and further research possibilities will be given.

2. Methodology: value-focused thinking

Value-focused thinking is a decision technique suggested by Keeney (1994). The approach calls for the identification of the stakeholders that will be impacted by the decision. These persons or groups of persons are then questioned about their values, concerning the specific area under consideration. Values are those principles that one strives to and define all that one can care about in a specific situation (Keeney, 1994). Next, the identified values are converted into objectives. An objective is characterized by three features, namely: a decision context, an object and a direction of preference (Sheng et al., 2005). For example, statements (values) such as "users should not execute strange e-mail attachments as they may contain viruses" can be changed into the objective "minimize virus infections". The decision context is related to infections, the objects are viruses and the direction of preference is to have less virus infections. It is of course possible to derive more than one objective from a specific value statement, e.g. to maximize the responsible use of e-mail is another objective that can be derived from the above value statement. Keeney states that one of the greatest benefits of this approach is that better alternatives for a decision problem can be generated once objectives have been established. This is in contrast with the more traditional method, called attribute-focused thinking, where alternatives are first identified after which the objectives are specified.

Following the determination of objectives, a process to distinguish between means and fundamental objectives is performed. Fundamental objectives refer to the objectives underlying the essential reasons for the problem being under consideration while means objectives are regarded as those whose attainment will help achieve the fundamental objectives (Nah et al., 2005). To perform this step, Keeney's "why is this important" test can be done. Each objective is evaluated against this question and if an objective is found to be important because it helps achieve another objective, it is categorized as a means objective; otherwise it is a fundamental objective. For example, an objective such as "minimize virus infections" is important as it helps achieve another objective "maximize integrity of data". It will therefore be classified as a means objective while the data integrity objective is a fundamental objective – it is regarded as an essential reason for the project and is not used to attain any other goal. Finally a means-ends objective network is constructed to show the interrelationships among all objectives. The network is then used to derive cause–effect relationships and to generate potential decision opportunities. The complete process and steps to be followed are shown in Fig. 1.

The value-focused thinking approach has already been applied successfully in different areas. Hassan (2004) applied it to the environmental selection of wall structures while Nah et al. (2005), used the approach to describe the value of mobile applications. Other examples can be found in the works of Dhillon and Torkzadeh (2001) and Dhillon et al. (2002) where the value-focused thinking approach was used in assessment of IS security in organizations and privacy concerns for Internet commerce. In this study the value-focused approach was applied at a university to identify key areas of concern to ICT security awareness.

3. Application of value-focused thinking

The value-focused thinking methodology was applied in a university environment and resulted in a network of means and fundamental objectives that highlighted key areas to be addressed in an ICT security awareness program for a university.

Modern universities are managed and operated along the same lines as any other corporation. However, the nature of universities implies that there may be different requirements and needs that necessitate adequate provision of ICT resources, safeguarding of information assets and optimum use of facilities.

At universities, academic freedom often influences the type of security controls that can be enforced. The result is that the enforcement of rigid security measures and controls are often more problematic than in other corporate environments. Another aspect is the fact that fixed academic programs are being followed and students from different disciplines and departments are often in the same classes. Facilities such as class rooms, laboratories, IT resources etc. are used according to schedules with a great interdependence among staff, students and the facilities – if any of these components in the university system is not available when required it will be very difficult, if not impossible, to re-schedule activities.

Some other examples on how universities, with their core business centered on teaching and research, may differ from profit making organizations include the following. In certain subjects, students usually do practical examinations and tests using desktop computers that are connected to a network – unavailability of any of these resources may lead to examinations not completed on time as well as possible loss of work already completed. Networks are increasingly used for presentation of study material and communication to students. Having said this, it is clear that, not just availability, but areas such as confidentiality, integrity, proper use and management of resources, etc. would also play a major role in examinations. Laboratory and research experiments may be time dependent in which case processing of data needs to be done immediately or at least within a certain time frame. In other cases, long and complex computations that take hours or even days to complete require uninterrupted availability of computing resources, a high level of precision, correctness and confidentiality. The effective use of resources will be seriously impacted in cases where work has to be redone. Lastly, just as any other business, universities rely heavily on ICT resources for their day-to-day operations. Their databases need to be as accurate and confidential as any other organization and resources should not be available when needed, it means that universities will not be able to effectively conduct their

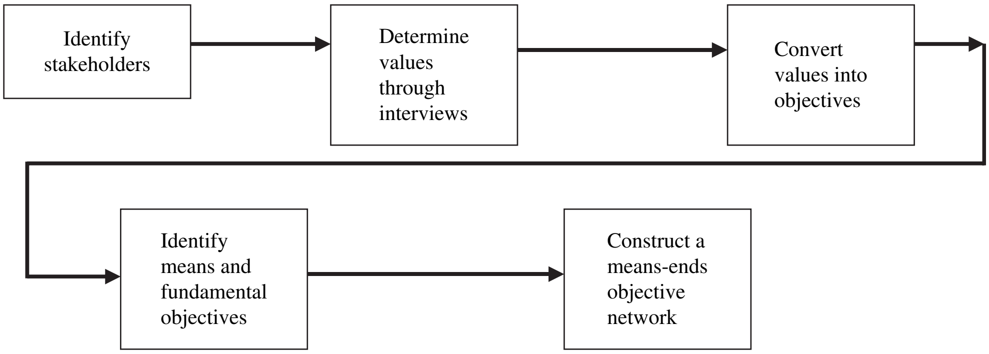

Identify stakeholders
Determine values through interviews
Convert values into objectives
Identify means and fundamental objectives
Construct a means-ends objective network

Fig. 1 – Value-focused thinking process.

business, productivity will be down and ultimately the image of the institution may be seriously affected.

Keeney's value-focused approach was used to conduct interviews and to organize the data into the required network. The primary objective of the interview process was to identify stakeholders' wishes, concerns, problems and values pertaining to ICT security awareness.

A discussion document, rather than a questionnaire, was used to obtain information from the interviewees. The discussion document contained six statements or questions and was compiled according to the techniques for the identification of objectives suggested by Keeney. The issues discussed with interviewees include:

1. What is important to you regarding ICT security awareness?
   The purpose of this discussion point was mainly to be able to identify goals and guidelines and to determine strategic and generic objectives. Typical answers such as the sensitivity and confidentiality of data, image of the institution, effective and efficient use of resources, etc. were given.
2. What would you do or implement to increase the level of ICT security awareness?
   The development of a wish list and identification of alternatives are important when trying to establish values that could ultimately lead to objectives. This second discussion point assisted greatly in this regard and to further encourage interviewees the statement ''if you have an unlimited budget'' was added to the question. The wish lists included many different answers such as installation of security cameras, provision of adequate backup facilities, proper distribution of information, etc.
3. What is your current concerns regarding ICT security awareness levels?
   The identification of shortcomings and problems helped in the process of describing objectives, e.g. a concern such as ''people write their passwords on pieces of paper'' indicated that the confidentiality of passwords needs to be addressed.
4. What can be done to raise and maintain ICT security awareness levels?
   This point is closely related to the second one but the goal here is to identify those aspects that can be fixed now, e.g. ''make the existing security policies more accessible'' as opposed to a wish list.
5. What are the limitations to raise ICT security awareness levels?
   In some cases the responses overlap with those obtained under the third discussion point. Limitations given by respondents include issues such as lack of interest, lack of management involvement and commitment, lack of resources, etc.
6. If you have to evaluate ICT security awareness levels, how would you do it and how would you know that the awareness level is acceptable?
   This question was added in an effort to quantify objectives. Answers ranged from statistical calculations to comparisons with similar institutions to the use of external sources such as consultants and research companies.

A similar interview process as the one used by Nah et al. (2005) was followed. In their study on the value of mobile applications, they interviewed key personnel, one at a time, and carried on interviewing employees until no new values or objectives could be identified. The 'saturation point' was reached after the seventh interview but a total of 10 employees were interviewed. This is an acceptable approach, in qualitative research, to determine a stopping point in the data collection process and the question of how long a researcher must continue to gather data is answered by Glaser and Strauss (1967). They use the term theoretical saturation which means no additional data is found by the researcher for a specific category in a study.

In this study, a total of seven employees were interviewed, however, no new values were obtained after the fourth interview. Statistically speaking, it might be argued that the use of only seven employees cannot be regarded as a sufficient sample. While it is also true that one would never know if the next employee would be able to provide new information, it was decided to keep to the generally accepted qualitative procedure utilizing the 'saturation point' stopping rule when collecting data. The number of employees interviewed (seven) also correlates with the number of interviews (10) in the Nah et al.'s (2005) study.

The interviews were recorded for future reference. Each interview lasted approximately one and a half hours. Respondents included staff from both management and non-management levels and were selected from the IT department and from users. The immediate result of the interview process was a list of values that apply to ICT security awareness which was converted into objectives following the process described in Section 2. The fundamental and means objectives were then derived from the list of objectives using Keeney's 'why is it important?' test. Finally the means-ends objective network was constructed graphically by linking means and fundamental objectives to one another to show the interrelationships among them. A more detailed discussion on this network follows in the next section.

4. Results

A network of objectives was constructed from the data obtained during the interviews and is presented in Fig. 2. On the left are the means objectives that show the concerns, wishes and values of the interviewees pertaining to ICT security awareness. The right hand side shows the fundamental objectives that are derived from the means objectives or stated by the stakeholders. For explanatory purposes, a description of how the objectives "maximize physical access control" and "maximize availability of data and hardware" were identified and classified as means or fundamental objectives, in Fig. 2, is given below.

First, the output from the interviews resulted in a list of value statements such as the following examples:

– Equipment should always be locked away
– Card access is important
– Access gates/turnstiles should always be in working order

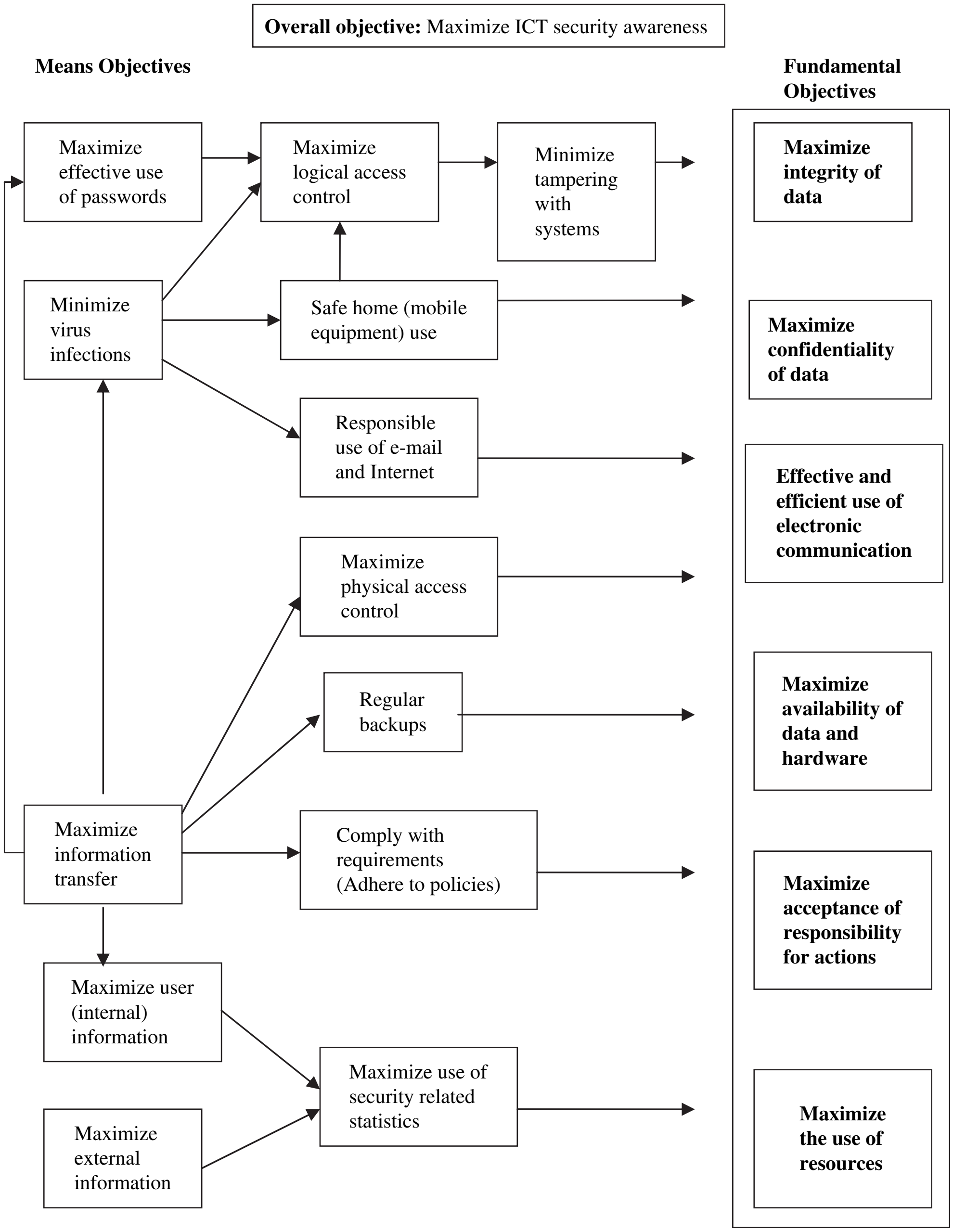

Overall objective: Maximize ICT security awareness
Means Objectives
Fundamental Objectives
Maximize effective use of passwords
Maximize logical access control
Minimize tampering with systems
Minimize virus infections
Safe home (mobile equipment) use
Responsible use of e-mail and Internet
Maximize physical access control
Regular backups
Maximize information transfer
Comply with requirements (Adhere to policies)
Maximize user (internal) information
Maximize external information
Maximize use of security related statistics
Maximize integrity of data
Maximize confidentiality of data
Effective and efficient use of electronic communication
Maximize availability of data and hardware
Maximize acceptance of responsibility for actions
Maximize the use of resources

Fig. 2 – Means-ends objectives network for ICT security awareness.

– Laptops should be secured with cables
– Security cameras should be installed
– Etc.

After careful consideration of these value statements and the reasons why respondents put them on the list, it became clear that physical security is regarded as an issue and that physical access control should be one of the objectives. This set of relevant value statements was then converted into the objective ‘‘maximize physical access control’’.

Other value statements such as

– Data should be available when it is needed
– Hardware and equipment should be available and in working condition
– Data should be backed up regularly and
– Data should be stored on network drives to ensure regular backups are taken

lead to the formulation of the objective ‘‘maximize availability of data and hardware’’. To classify these objectives, the ‘‘why is this important?’’ test, discussed in Section 2, was used. It is clear that the physical access objective is important in order to satisfy other needs such as availability of data and hardware, integrity of data, effective use of resources, etc. It was therefore classified as a means objective. The other objective ‘‘maximize availability of data and hardware’’ was not necessary to achieve any other objective. It was seen as one of the essential reasons why ICT security awareness should be addressed and was therefore classified as a fundamental objective.

The fundamental objectives, found in this study, are in line with the acknowledged goals of ICT security, e.g. integrity, confidentiality and availability (BS7799, 1999). Other objectives that emerged from this exercise are more on the social and management side, e.g. responsibility for actions and effective use of resources. Although no new aspects of security awareness could be identified, the approach served as confirmation that the same areas, found in any corporate environment, is important in an academic environment and could not be ignored. In addition, the results are important as the information will be used to develop a measuring instrument that will cover and focus on the identified means objectives in order to address fundamental objectives. It may also serve as a framework for management to structure an awareness program that includes all the appropriate learning areas.

The fundamental and means objectives derived from the network are listed in Tables 1 and 2. Table 2 can be used to see what aspects influence the means objectives according to the interviewees while Table 1 shows the fundamental objectives and the factors describing them.

Table 1 – Fundamental objectives

1. Maximize integrity of data
   - Correctness of data; comply with formal ICT strategy
2. Maximize confidentiality of data
   - Ensure confidentiality of business, research, client (student) data
3. Effective and efficient use of e-communication systems
   - Minimize cost of e-communication; maximize e-communication resources
   - Minimize negative impact of e-communication
4. Maximize availability of hardware and software
   - Uninterrupted usage; prevent damage and loss
5. Maximize acceptance of responsibility for actions
   - Consequences of actions
6. Maximize the use of resources
   - Comply with formal ICT strategy

4.1. *Fundamental objectives*

4.1.1. *Maximize integrity of data*

Integrity can be defined as the need to ensure that information has not been changed accidentally or deliberately and that it is accurate and complete (SANS, 2006). Several potential threats exist in a university environment that can result in the corruption or loss of data and information. The loss of data can be a result of hardware failure (availability) or can be attributable to user actions. Accidental errors by users such as data capturing mistakes or erasing files or disks may occur, but there is also deliberate user actions that may compromise data integrity, e.g. change of test and examination marks, tampering with business and research results, etc. A third potential threat to the loss of data can occur as a result of system contamination often caused by computer viruses. Some of the aspects that emerged from the value-focused assessment exercise, relating to the threats, include the maximization of both physical and logical access, safe use of mobile equipment and responsible handling of e-mail and Internet facilities.

4.1.2. *Maximize confidentiality of data*

Confidentiality is the ability to avoid disclosing information to anyone who is not authorized to use it (Pipkin, 2000). In a university environment the need for nondisclosure is perhaps not so high in comparison with other profit making organizations. However, the need for confidentiality exists and the protection of personal data (privacy) for students and staff, and the protection of data belonging to the university (secrecy) are important in applications such as student systems, financial systems, research projects, employee systems and others. Tampering with systems (physical and logical access) was highlighted as an important area by stakeholders at the university under review.

4.1.3. *Effective and efficient use of electronic communication resources*

The Internet is a global network that connects millions of computers all over the world and enables users to exchange data, news, opinions, etc. A system of Internet servers forms the World Wide Web which supports specially formatted documents, links to other documents, graphics, video and audio files. An important service offered by the Internet is the transmission of messages or the use of e-mail. At universities the use of the Internet and e-mail is an essential resource for both academic and non-academic staff. Electronic business and commerce systems are common at universities, marketing information is made available to prospective students

Table 2 – Means objectives

1. Maximize effective use of passwords
   - Use of strong passwords, keep passwords secret; sign off from PC when not in use
   - Minimize number of passwords used on the Internet
2. Maximize logical access control
   - Use encryption; limit multiple log-ins; correct system authorizations
   - Separation of duties, clean-up procedures when people resign
3. Minimize tampering with systems
   - Restricted access
4. Minimize virus infections
   - Viruses from home/mobile equipment; viruses from Internet; viruses from e-mail
5. Safe home (mobile equipment) use
   - Remote access/modems; use of passwords
6. Responsible use of e-mail and Internet
   - Cost and time for Internet usage; handle strange e-mails with care: large attachments
   - Correct defaults
7. Maximize physical access control
   - Minimize theft; use of security equipment, e.g. cameras; clean-up procedures when people resign
8. Make regular backups
   - Minimize loss of data; criteria on how long to keep data
   - Correct default saves; criteria for important data; availability of equipment to make backups
9. Maximize information transfer to employees
   - Maximize IT literacy; use communication channels (posters, bulletin boards, contracts)
   - Criteria to identify important data; illegal use of software
10. Comply with requirements (Adhere to policies)
    - Maximize IT literacy; use communication channels (posters, bulletin boards, contracts)
    - Make risks clear; make security implications clear
11. Maximize user (internal) information
    - Use user feedback; use internal audit statistics; minimize loss of knowledge, e.g. when people resign
12. Maximize external information
    - Use external input/reports, e.g. external auditors, Gartner
13. Maximize use of security related statistics
    - Use all comparable statistics

through web pages and in a global research environment, these facilities has become indispensable. Electronic communication is also used in the teaching function – both as a subject of study as well as a tool and an aid to perform teaching activities. Unfortunately, the Internet and e-mail facilities expose organizations and universities to a variety of risks. Some of the more common risks include spreading of viruses (the most important threat identified in this case study), using the facilities to conduct private business, routing chain e-mails, impersonation, eavesdropping, etc. Examples of the types of losses that can be incurred are damage to data, systems and networks, monetary damage, compromised privacy, reputation damage through changed or lost information, denied services, etc.

4.1.4. *Maximize availability of data and hardware*

Availability, which is often associated with service level agreements, can be defined or classified in various ways, e.g. in terms of data availability, system availability, application availability, infrastructure availability, etc. For the purpose of this study the following general definition is used (Pipkin, 2000). Availability is the state of being able to ensure that users can use any information resource whenever and wherever it is needed in accordance with applicable privileges. For the university under review in this study, two main aspects that have an impact on availability were identified. First, the making of backups was pointed out as an area for improvement. Most staff does not make regular backups of their work. There is an apparent lack of understanding on what to backup, where and how regularly backups should be made. Staff is encouraged to save their data, documents, projects, etc. on a network drive in order for it to be backed up automatically but apparently this is not done. Secondly, the implementation and enforcement of proper physical access controls would assist in preventing damage to and loss of ICT resources.

4.1.5. *Maximize acceptance of responsibility for actions*

There are a variety of groups and individuals that are responsible for different activities regarding security issues. In general, for this study, the assumption is that users should be held responsible for their actions. According to the Concise Oxford English Dictionary (2002) the term responsible is defined as ''being the primary cause of something and so able to be blamed or credited for it; answerable to; morally accountable for one's behaviour; capable of being trusted''.

Responsibility is then simply described as ''the state or fact of being responsible''. It means that if users at the university do not comply with the university's requirements and policies, e.g. not making regular backups or not installing anti-virus software on their PC's, etc., they will be held accountable for any negative results such as loss incurred, degradation of system performance or loss of confidence from clients. To assist in establishing a culture where users will accept that they are responsible for their actions, the areas of information transfer and security policies were identified as key issues.

4.1.6. *Maximize the use of resources*

Information technology resources, which include money, time and people, are made available to support a university's scholarly, educational, research and administrative activities. These resources are limited and should be utilized in a well managed way. Security transgressions may have a direct impact on cost (e.g. as a result of loss of data, equipment or client confidence), time (e.g. to attend to problems) and staff (e.g. to do recovery work). The areas of information transfer and adhering to security policies were once again identified in this study as areas to be focused on to ensure the maximization of resources.

5. Conclusion and further research

The overall objective that resulted from this study is the maximization of ICT security awareness to aid the university in providing a sustainable service to its staff, students and other stakeholders. Fundamental objectives to achieve this were identified by constructing a network using the value-focused approach. These objectives can serve as a basis for decision making and to guide the planning, shaping and development of ICT security awareness programmes and ultimately to influence the general information security culture in a company. ICT security awareness programmes can be used to train staff and sensitize them in the security arena to get a more secure environment compliant to standards such as BS7799 (1999) and others. Information security culture, which forms part of organizational culture, has to do with employees' behavior. The way things are done is based on collective values, norms and knowledge and will have a crucial impact on corporate success (Schlienger and Teufel, 2003). To instill an information security culture is not an easy task and Schlienger and Teufel (2003) suggested that the process and management thereof should be treated as a never-ending cycle of evaluation and change or maintenance. Key to the evaluation phase is the identification of ''what to analyze?'' The result from the study described in this paper greatly assists in identifying and describing those areas that need to be included in an information security awareness and culture program. Furthermore, as security culture is so closely related to security behavior, it is believed that the analyzing of security awareness levels will directly contribute to the establishment and maintenance of a security culture.

The intention is to perform a follow-up study and some of the goals would be to generate measures representing the factors of the identified focus areas, and the development of a final model, using appropriate management science techniques, to generate measurements and recommendations that are reliable and valid.

Acknowledgement

The authors would like to thank the two anonymous referees for the useful comments and suggestions made. The authors alone are responsible for any errors and omissions.

REFERENCES

BS7799. Code of practice for information security management. UK: British Standards Institute; 1999.

Czernowalow M. Lack of policy causes IT risks. Available from: ITWEB, <http://www.itweb.co.za>; 15 July 2005.

Dhillon G, Torkzadeh G. Value-focused assessment of information system security in organizations. In: Proceedings of the 22nd international conference on information systems; 2001. p. 561–6.

Dhillon G, Bardacino J, Hackney R. Value focused assessment of individual privacy concerns for Internet commerce. In: Proceedings of the 23rd international conference on information systems; 2002. p. 705–9.

Glaser BG, Strauss AL. The discovery of grounded theory: strategies for qualitative research. New York: Aldine de Gruyter; 1967.

Hassan OAB. Application of value-focused thinking on the environmental selection of wall structures. Journal of Environmental Management 2004;70:181–7.

Keeney RL. Creativity in decision making with value-focused thinking. Sloan Management Review 1994;Summer:33–41.

Nah FF, Siau K, Sheng H. The value of mobile applications: a utility company study. Communications of the ACM 2005; 48(2):85–90.

Pearsall J, editor. Concise Oxford English Dictionary. 10th ed. Oxford University Press; 2002.

Pfleeger CP, Pfleeger SL. Security in computing. 3rd ed. Prentice Hall; 2003.

Pipkin DL. Information security: protecting the global enterprise. Prentice Hall; 2000.

SANS. Sysadmin, Audit, Network Security Institute; 2006. <http://www.sans.org/resources/glossary.php#> [accessed 1 February 2006].

Schlienger T, Teufel S. Information security culture – from analysis to change. South African Computer Journal 2003;31: 46–52.

Sheng H, Nah FF, Siau K. Strategic implications of mobile technology: a case study in using value-focused thinking. Journal of Strategic Information Systems 2005;14(3):269–90.

Whitman ME, Mattord HJ. Principles of information security. 2nd ed. Thomson; 2005.

L. Drevin is a lecturer in Computer Science and Information Systems, at the North-West University (Potchefstroom Campus) in South Africa. She joined the university in 1985 as a staff member. She has an M.Sc. in Computer Science and Information Systems. Her current interests include security awareness and education, and project failures.

H.A. Kruger is an Associate Professor in the School of Computer, Statistical and Mathematical Sciences at the

North-West University (Potchefstroom Campus) in South Africa. He previously worked for Anglo American Corporation as a senior Computer Auditor and has more than 10 years experience in Information Risk Management. He has a Ph.D. in Computer Science, an M.Com. in Information Systems and an M.Sc. in Mathematical Statistics. His current interests include decision modeling, security awareness and the use of linear programming models.

T. Steyn is a Professor in Computer Science in the School of Computer, Statistical and Mathematical Sciences at the North-West University (Potchefstroom Campus) in South Africa. He joined the university in 1974 as a staff member of the then IT department and moved to Computer Science in 1980. He has a Ph.D. in Computer Science. His current interests include cutting stock optimization, security awareness and databases.

ELSEVIER

Computers & Security

Bridging the gap between general management and technicians – A case study on ICT security in a developing country☆

Jabiri Kuwe Bakari, Charles N. Tarimo, Louise Yngström, Christer Magnusson, Stewart Kowalski*

Department of Computer and System Sciences, Stockholm University/Royal Institute of Technology, Forum 100, SE-164 40 Kista, Sweden

Keywords:
ICT security management
Organisations
Management
Technicians
Perception problem
ICT-related risks

ABSTRACT

The lack of planning, business re-engineering, and coordination in the whole process of computerisation is the most pronounced problem facing organisations. These problems often lead to a discontinuous link between technology and the business processes. As a result, the introduced technology poses some critical risks for the organisations due, in part, to different perceptions of the management and technical staffs in viewing the ICT security problem. This paper discusses a practical experience on bridging the gap between the general management and ICT technicians.

1. Introduction

The paper outlines a successful mission of how to bridge the gap between general management and ICT technicians. It is based on practical experiences obtained from an ongoing study which aims at developing guidelines for managing ICT security in organisations generally. The study was initially conducted in mid-2004 in five organisations in Tanzania in order to make preliminary observations. Later, at the beginning of 2005, one organisation was earmarked as a test-bed for further observations and here we present some of the findings. The organisation is a government-based service provider operating in 21 out of 26 regions in the country. The organisation has 900 staffs in total and its operations are based on four core services, where three of them are greatly dependent on ICT to meet their intended objectives. The organisation has approximately 2 million customers scattered throughout the country with approximately 25% active customers.

The study was guided by using the Business Requirements on Information Technology Security (BRITS) framework where risks are viewed as part of the actual business rather than primarily as part of the ICT, used together with the Security by Consensus (SBC) model where ICT security is viewed as a social technical problem (Kowalski, 1994; Magnusson, 1999; Bakari, 2005). BRITS is a systemic-holistic framework, combining finance, risk transfer, IT and security in a coherent system. The framework attempts to bridge the gap between top management and IT personnel by translating the financial language into the IT and IT security languages, and vice versa. The translation is achieved by making use of a repository of

☆ The paper was originally presented and published in the Security Culture Workshop at the IFIP SEC2006. This is an updated version.

* *Corresponding author*. Tel.: +46 08 674 72 37; fax: +46 08 703 90 25.
E-mail addresses: si-jba@dsv.su.se (J.K. Bakari), si-cnt@dsv.su.se (C.N. Tarimo), louise@dsv.su.se (L. Yngström), cmagnus@dsv.su.se (C. Magnusson), stewart.kowalski@ericsson.com (S. Kowalski).

mitigation suggestions, hosted in the Estimated Maximum IT Loss (EMitL) database (Magnusson, 1999; Bakari et al., 2005a,b). In the study the SBC model was used to view and explain security problems as layers of social and technical measures. In addition to these two methods, we also at different stages of the study made use of other methods, namely Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE), ITIL (IT Infrastructure Library), Control Objectives for Information and related Technology (COBIT) and the internationally recognised generic information security standard, comprised of a code of practice and a specification for an information security management system (ISO 17799). OCTAVE is a risk-based strategic assessment and planning technique for ICT security (Alberts and Dorofee, 2003). ITIL is a framework for IT management and COBIT is an IT governance framework and supporting toolset that allows managers to bridge the gap between control requirements, technical issues and business risks (ITIL, 2005; ISACA, 2005; ISO 17799).

The findings are organised in a list of 10 initial steps or aspects of importance to successfully bridge the gap. The presentation highlights the motivation and practical experiences of each step.

2. The ten aspects of importance in bridging the gap between the management and technicians

In this section, the 10 steps are introduced and later the experience encountered when executing each step is presented. The steps include:

(i) Getting top management's backing (the chief executive officer (CEO) buying into the idea first)
(ii) Getting technical management backing (the technical department is the custodian of ICT in an organisation)
(iii) Setting up the special ICT security project team (start by forming a provisional ICT security task force)
(iv) Quick scan of the ICT-related risks and their consequences for the organisation (risk exposure due to ICT)
(v) Getting management's attention and backing (the management as a whole need to buy into the idea as well)
(vi) Getting the current status of ICT security documented (take stock of the existing situation)
(vii) Conducting awareness-raising sessions (to allow staffs to recognise ICT security problems and respond accordingly)
(viii) Carrying out risk assessment/analysis
(ix) Working out the mitigation plan (short-term plan for issues that need immediate attention and long-term plan)
(x) Developing countermeasures

2.1. *Step 1: getting top management's backing (the CEO buying into the idea first)*

ICT security appeared to be a new concept to most CEOs in the organisations studied. As confirmed by numerous researches, management backing is important in any effort to improve security in organisations as suggested in some studies (Alberts and Dorofee, 2003; Caralli, 2004; Solms and Solms, 2004; Solms, 2005) and also appears as an important factor in corporate governance as discussed in Control Objectives for Information and Related Technologies (COBIT) by Information Systems Audit and Control Association (ISACA). However, getting an appointment to meet the CEO and talk about ICT security was not easy. In most cases, we were directed to see the IT director or chief security officer. Here one needs to be patient and accept a long-awaited appointment to see the CEO who is always busy and in this case the time slot would be only 10–15 min. Nevertheless, we achieved our goal of meeting them and introduced our agenda on what ICT-related risk is all about and what the challenges are in managing such types of risks. Further, the consequences of not managing such risks for the shareholder value were also discussed, emphasising that today's CEOs will be responsible to their board for the state of ICT security in their organisations. All these were discussed with respect to risk exposure to key performance indicators which may affect the organisation from reaching its mission and business objectives. An example of risk exposure discussed was the problem of business interruption which can be propagated through to the balance sheet with great financial implications and cause embarrassing media coverage, loss of confidence by customers and staffs, resulting in loss of credibility.

2.2. *Step 2: getting technical management backing (technical departments are the custodians of ICT in an organisation)*

It was hard and in most cases almost impossible to talk about ICT-related issues in the organisation without the permission of its IT department. This was mainly due to a perception problem which is also discussed in Bakari et al. (2005a) where the complex problem of ICT security has been relegated to the IT department or rather treated as a technical problem, with no relevant and operational organisation-wide ICT security policy and procedures. Most of those we asked for an appointment gave the following reaction: ''Have you consulted the IT department?'' On the other side, the technical staffs are aware of the ICT security problems, though mostly as a technical concern and not as a business concern. In order to get their support, we had to describe the security problem more holistically, i.e. including both technical and non-technical issues and the reasons why we should include and talk to other departments as well. Our observation indicated that the difference in perception between the management and the technical department made it difficult for the technical department to address the problem adequately. An attempt was made by Bakari et al. (2005b) to address this perception problem using EMitL tool as suggested in the BRITS framework by Magnusson (1999). Getting technical staffs to understand the non-technical components of the problem and how to communicate the problem to the management as risk exposures which needed its attention was yet another important step to take. There were concerns from senior technical staffs on how we were to make the management understand the problem, and what language to use for them to understand.

This was asked by senior staffs, when we were preparing to meet the management team in one of the talks (step 5).

2.3. Step 3: address the ICT security problem as a special project (forming a provisional ICT security task force)

> *"Organisations can no longer be effective in managing security from the technical sidelines. Security lives in an organisational and operational context, and thus cannot be managed effectively as a stand-alone discipline. Because security is a business problem, the organisation must activate, coordinate, deploy, and direct many of its existing core competencies to work together to provide effective solutions."*
>
> *(Caralli, 2004)*

After succeeding in getting the support of the top management and technical management, the important question at this stage was how or where do we start? It was at this stage that we formed the special ICT security project team. The composition of the team included three technical staffs (software, network and hardware), one legal officer, one internal auditor, one security (physical/traditional) officer, and one member of staffs from operational departments (where core services of the organisation are processed). Also one more member of staffs from the insurance department was in the team purposely for risk management as there was no department other than insurance to handle/manage risks in the organisation. All team members were senior members of staffs who have worked with the organisation for more than five years. The main question we faced here was then why to choose staffs from these departments? Our response was based on the facts below and which are also similar to OCTAVE (Alberts and Dorofee, 2003) principles, where the interdisciplinary ICT security project team is staffed by personnel from the organisation itself:

Technical: partly the ICT security problem is a technical issue which could be a result of software, hardware or network problems. In the case of software, there are mostly two fundamental problems, one that affects the application system software and the other that affects the Operating System software. Both could be as a result of the introduction of malicious software (virus, worms, etc.) or system failure, either due to power failure or some other reasons. In the case of hardware, there could be physical damage, a power problem or the hardware or part of it being stolen. Network security can be a combination of both; problems that affect the software and hardware parts of the network. Technical staffs who are working in these sections would be in a good position to give more information about their experience regarding ICT security from a technical point of view.

Auditors: traditionally, auditors are used to audit the financial transactions or operational processes and compliances to laws and regulations, policies, standards and procedures. Given the nature of their work they can also stand back and see the big picture concerning the risk exposure facing an organisation. Auditing ICT is usually considered operational. The one prime focus for ICT audit is security – evaluating whether the confidentiality, integrity and availability of data and services are ensured through the implementation of various controls (ISACA, 2005). It also involves evaluating the realisation of benefits to the business from its investment in IT. Apart from building capacity for internal ICT audit, including the transition from traditional auditing to a hybrid type of auditing (meaning the auditing includes information systems), informed corporate auditors can be in a better position to provide the information needed to advise the management on the importance of paying more attention to ICT security management than technicians' advice.

Legal: as the dependency on ICT in an organisation grows, legal issues such as liabilities, in particular computer/cyber crime (a violation of the law committed with the aid of, or directly involving, a computer or data processing system) are becoming an indispensable part of ICT risk management. Involvement of a legal officer in the team facilitates in addressing the ICT security problems from a legal perspective.

Security: most of the security departments – particularly in the studied organisations – still value physical assets, which means that security strategies end up taking more care of tangible assets than intangible ones. For example, currently CCTVs (close-circuit TVs) are installed in the reception area and along the corridors but not in the server rooms which keep valuable information assets. This situation as revealed here discourages one from stealing tangible company assets from the building as there is a chance of being seen. However, for someone who aspires to steal *information assets*, will have a free ride. By not having the server rooms monitored properly – apart from those who can compromise the assets through the network – it is implied that the security monitoring system is meant for outsiders. Thus, the involvement of physical security staffs helps to identify what needs to be re-engineered for the existing security in the organisation.

Operations: operations is where the core services of the organisation take place. Thus, operations can be an area where the greatest adverse impact on the organisation's mission and business objectives can be observed. In our work we considered having a senior member of staffs from the operations department who is fully knowledgeable of operational transactions. His participation in the team assists in highlighting the operational processes and identifying critical areas of operation. This is an important component in the risk assessment exercise.

Insurance/risk manager: ICT security management is basically risk management focusing on ICT – mainly how to insure valuable assets, including ICT assets.

Human resources: human elements in security tend to be the weakest link in any security chain, even where the best technical security controls are in place (Bishop, 2003). A simple social engineering action, which is a result of not ensuring that staffs are aware of the risks and are familiar with sensible and simple security practices, can ruin the organisation's credibility. Therefore, a strong ICT security management programme cannot be put in place without significant attention being given to human resources. People from the human resources department/unit are responsible for personnel security which, according to ISO 17799, covers not only permanent and temporary staffs of the organisation but also extends to contractors, consultants and other individuals working on the organisation's premises or using the organisation's information and information processing assets.

Furthermore, the human resources department is responsible for recruitment, terms and conditions of employment, including job descriptions and termination. It is also responsible for awareness-raising and training of staffs on security policy, procedures, and techniques, as well as the various management, operational and technical controls necessary and available to secure ICT resources, once such measures are in place (Wilson and Hash, 2003). Inclusion of a human resource person in the team from the beginning helps to take into consideration the human aspects of the security problem from the outset, when designing the ICT security management programme.

Finance: there are two main reasons why finance should be considered. First, all operations of the organisation depend on financial transactions. The roles and responsibilities of staffs vary as compared with other departments due to the nature of their work – financial transactions. Unclear roles and responsibilities can be tolerated in the manual system but not in the computerised financial information system. Secondly, in most cases the end result of any security incident has financial implications; sometimes damage can be propagated to the organisation's final balance sheet.

Selection criteria: generally, the selection was strictly of staffs who have spent a substantial amount of time in the area of consideration and who are also ICT literate, for example, senior auditor with some general computer knowledge. The team was subjected to several orientations about ICT security management in general with reference to the organisation.

2.4. Step 4: quick scan of the ICT-related risks and their consequences for the organisation (risk exposure due to ICT)

Before meeting the management as a whole, we needed some kind of justification or evidence of ICT-related risks and their consequences for the organisation. This was obtained by first working out some facts on the likely consequences of ICT-related risks for the organisation. We achieved this by carrying out a quick scan of such risks with the help of the ICT security team. This exercise involved capturing information on what the organisation is doing and how its core services are linked to the use of ICT and hence what kind of risk exposures and their consequences for the organisation. Face-to-face interviews with the CEO, chief financial officer (CFO), IT managers and the heads of the departments involved in the provision of the core services were conducted.

Our interview questions were based on OCTAVE processes 1 and 2, which are primarily for gathering information on the senior management's and operational area management's views of ICT assets, areas of concern, security requirements, current security practices and current organisational vulnerabilities. The two processes are among the four used in OCTAVE phase 1 when building asset-based threat profiles of an organisation as detailed in Alberts and Dorofee (2003, p. 46).

We used the collected information to figure out how the organisation's objectives are supported by ICT assets and in turn what are the possible risks to, and consequences for, the organisation's business objectives as shown in Fig. 1.

We also made use of EMitL tool in an attempt to translate what management sees as damage exposure to corresponding ICT-related risks and hence ICT security properties as shown in Fig. 2.

The tool helped to interpret the technical terminologies of the consequences of losses in the corporate value, based on financial indicators. This interpretation was based on three groups of damage exposure due to ICT risks, namely liability claims, direct loss of property and business or service interruption; also explained in the works by Bakari and co-workers (2005, 2005b). The damage exposures are in turn mapped to ICT security properties.

2.5. Step 5: getting management's attention and backing (the management as a whole buy into the idea as well)

The management had to be convinced and understand that their organisation was vulnerable to ICT-related risks. Furthermore, we had to educate them on the magnitude of the security problem, and insist that ICT security was more than technology and more of a human issue. This means it

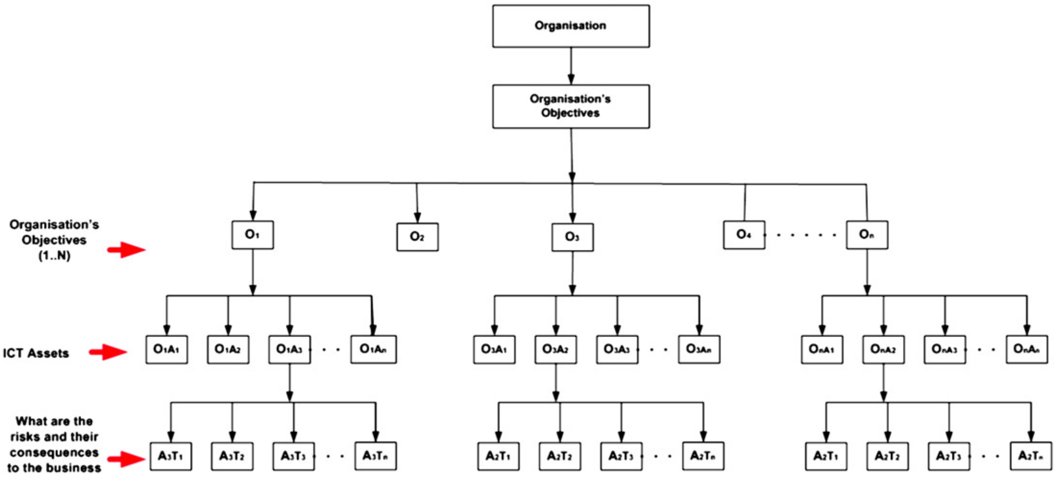
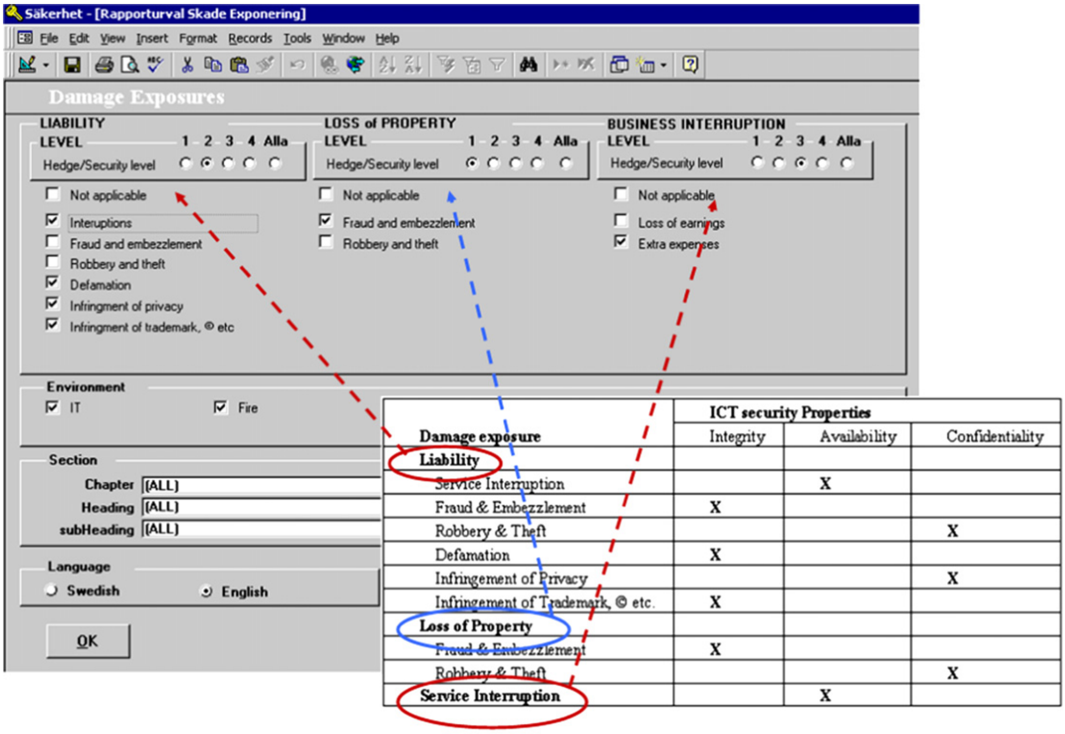


Fig. 1 – Deriving risks to, and consequences for, the organisation's business objectives.


Säkerhet - [Rapporturval Skade Exponering]
File Edit View Insert Format Records Tools Window Help

Damage Exposures

LIABILITY
LEVEL 1 2 3 4 Alla
Hedge/Security level ○ ● ○ ○ ○
□ Not applicable
☑ Interuptions
□ Fraud and embezzlement
□ Robbery and theft
☑ Defamation
☑ Infringment of privacy
☑ Infringment of trademark, © etc

LOSS of PROPERTY
LEVEL 1 2 3 4 Alla
Hedge/Security level ● ○ ○ ○ ○
□ Not applicable
☑ Fraud and embezzlement
□ Robbery and theft

BUSINESS INTERRUPTION
LEVEL 1 2 3 4 Alla
Hedge/Security level ○ ○ ● ○ ○
□ Not applicable
□ Loss of earnings
☑ Extra expenses

Environment
☑ IT ☑ Fire

Section
Chapter [ALL]
Heading [ALL]
subHeading [ALL]

Language
○ Swedish ● English

OK

Damage exposure	ICT security Properties		
	Integrity	Availability	Confidentiality
Liability			
Service Interruption		X	
Fraud & Embezzlement	X		
Robbery & Theft			X
Defamation	X		
Infringement of Privacy			X
Infringement of Trademark, © etc.	X		
Loss of Property			
Fraud & Embezzlement	X		
Robbery & Theft			X
Service Interruption		X	

Fig. 2 – Business damage exposure mapped to ICT security properties.

has something to do with the kind of policies and procedures that were in place in the organisation; the type of administration in place, the legal and contractual obligations the organisation had, particularly in delivering services, and also the ethics and culture of the individual staff member. The objective of this step was achieved by presenting to the management the worked-out status of their ICT security obtained in step 4. Diagrammatic representation of the problem and how ICT security was being addressed in their organisation helped to get their attention.

Fig. 3 shows how the problem was perceived on the left-hand side; and on the right-hand side of the figure can be

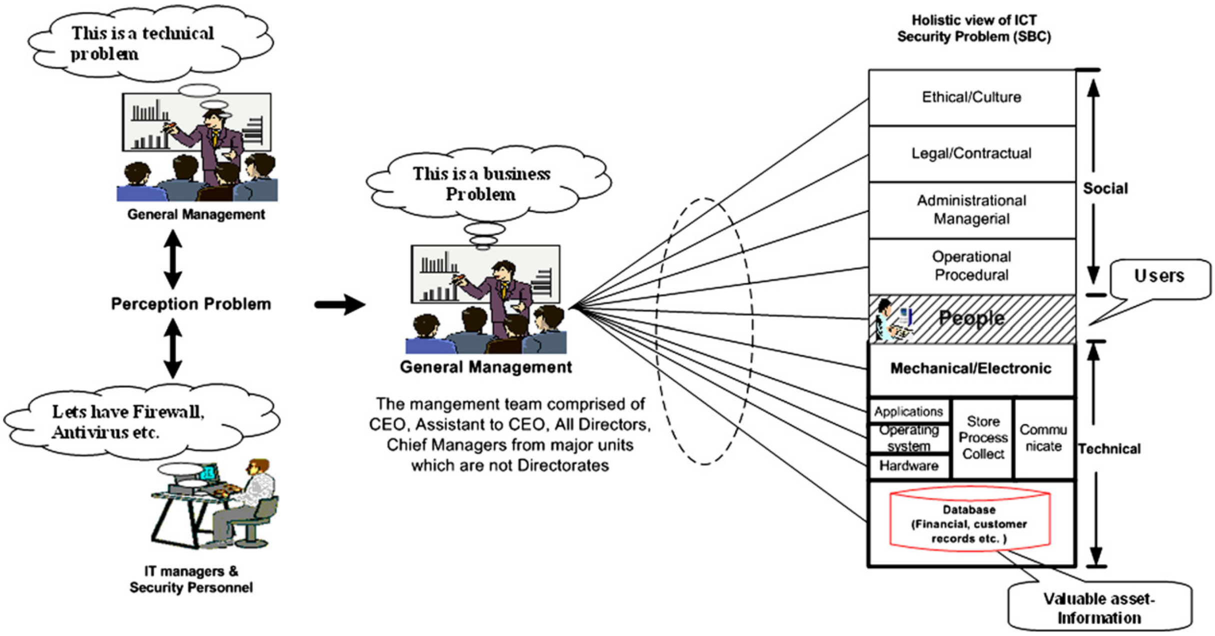

This is a technical problem
General Management
Perception Problem
Lets have Firewall, Antivirus etc.
IT managers & Security Personnel
This is a business Problem
General Management
The mangement team comprised of CEO, Assistant to CEO, All Directors, Chief Managers from major units which are not Directorates
Holistic view of ICT Security Problem (SBC)
Ethical/Culture
Legal/Contractual
Administrational Managerial
Operational Procedural
People
Mechanical/Electronic
Applications
Operating system
Hardware
Store Process Collect
Commu nicate
Database (Financial, customer records etc.)
Social
Users
Technical
Valuable asset-Information

Fig. 3 – How the ICT security problem was perceived and the way it was addressed.

seen a holistic view of the ICT security problem, with people sandwiched between the social and technical aspects, being an extension of SBC model (Kowalski, 1994). For example, we were able to show to the management team, which constituted the CEO, CFO, human resources manager, chief legal officer, chief security officer, chief internal auditor, operational manager, and planning and investment manager, technical managers and other managers, where and how their functions fit into the model. We also highlighted the problem in each dimension and discussed their role in managing the problem with respect to their positions in the organisation.

This was approximately a one-and-a-half hour session with the entire management team. The fact that ICT security management is a multidimensional discipline, as depicted in Fig. 3, was emphasised. We were able to convince the management that this is a business problem, with an ethical/culture dimension, an awareness dimension, a corporate governance dimension, an organisational dimension, a legal dimension, an insurance dimension, a personnel/human dimension, an audit dimension, and finally a technical dimension, also discussed in length in the study by Solms and Solms (2004). It is a socio-technical problem (Kowalski, 1994). We used the figure to demonstrate the management how ICT security is currently being managed in their organisation. The demonstration showed that, currently, the focus is mostly on the technical aspect, meaning that the existing countermeasures are mainly addressing the technical dimension which corresponds to the second and third signs of the 10 deadly sins of information security management as discussed in Solms and Solms (2004). Referring to ISO 17799 and COBIT for the organisational dimension, a management framework should be established to initiate the implementation of information security within the organisation. By using SBC framework, we were able to bring the management team together and discuss the security problem as a business problem as shown in Fig. 3. We were able to point out, with examples, what the security issues in each dimension are and the areas of consideration and consequences if such an issue is not addressed. For example, by highlighting to the management that ‘‘ensuring that staffs/users are aware of information security threats and their consequences for the organisation's mission and business objectives, in the course of their normal work’’ was the responsibility of people from the human resources department, helped them to see that this is not a technical department responsibility.

2.6. *Step 6: getting the current status of ICT security documented (take stock of the existing situation)*

Our next step was to have an idea of what existed in the organisation with respect to ICT security. This exercise involved taking stock of what existed in terms of: systems (hardware, software, platforms, networks, applications, users and assets); environment (location and services); security (threat types and potential ones and countermeasures that are currently in place); and procedures (any policies and procedures in place). This information helped to identify the current status of ICT assets, their location and the type of services they provide, as well as threat types for each identified asset and the security measures that are in place. We made use of OCTAVE phase 2, process 5 in some of the stages (Alberts and Dorofee, 2003, p. 49). The OCTAVE phase 2 deals with identification of key components of an organisation's information system. This exercise gave the security team more knowledge of ICT assets, their link to the organisation's objectives and helped to highlight areas that needed immediate attention. In addition, we later used this information during the awareness-raising sessions to help staffs understand and appreciate the types of ICT security problems they have. For example, we established that most of the different types of operating systems currently in use have been un-patched since they were bought; some have security features which are not enabled, and some have no security features at all; the licence status is not clear concerning some of the software; and the existing policy was not helpful as it was outdated and only a few senior staffs knew of its existence.

2.7. *Step 7: conduct awareness-raising sessions among users (with some feedback from steps 1–6)*

At this moment we had gathered information about the organisation, information systems risks and their consequences. We had the full support of the management and the now well-informed internal security team. It was at this step that we rolled out the awareness-raising sessions in the organisation. Our approach was top down as shown in Fig. 4. We

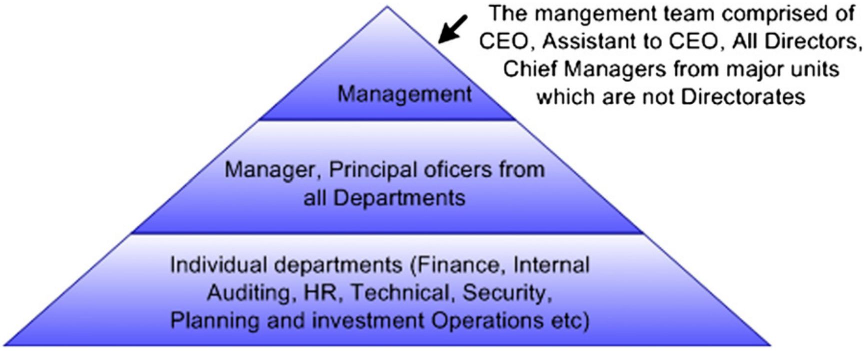

The mangement team comprised of
CEO, Assistant to CEO, All Directors,
Chief Managers from major units
which are not Directorates
Management
Manager, Principal oficers from
all Departments
Individual departments (Finance, Internal
Auditing, HR, Technical, Security,
Planning and investment Operations etc)

Fig. 4 – Awareness-raising sessions plan.

started with the management and the topic was ''Managing Technology risks, the role of the management, which included legal issues in a computerised environment''.

Along with the presentation notes, we attached the timetable of other training sessions for their departments/staffs as well. This helped to get the message across to other staffs through their bosses who made sure that their staffs attended their respective sessions. The second group was comprised of managers and Principal Officers from all departments. A similar topic was delivered but with a different emphasis from the one used with the management – strategic level. Here the emphasis was mainly on tactical and operational issues. More than 90% of the target groups attended the awareness-raising sessions in person. We made some observations during the sessions. For example, as we looked at the faces of staffs as they were arriving at the awareness-raising session room, we could read their faces saying, ''This session is not for me''. However, after some time into the session the situation changed and people were getting concerned about the issues being discussed. ICT security awareness-raising efforts were designed to allow staffs from various departments to recognise ICT security concerns, participate effectively in the ICT security management process and respond accordingly as suggested in the study by Wilson and Hash (2003), where detailed discussion on 'Building an Information Technology Security Awareness and Training Program' is presented. Apart from the general awareness-raising session, we also had special sessions with individual departments, namely legal, accounts, internal auditing, physical security, human resources and technical. For each session the focus was in accordance with their respective speciality.

For the Legal section, for example, the main point of discussion was what the ICT risks are from the legal perspective and hence the legal issues in a computerised environment. Some questions were posed such as: what are the implications of using unlicensed software in the organisation? How could a crime committed through computers be handled? Are the cooperate lawyers conversant with the subject matter? If not, what are the implications for the organisation should such a problem arise? What we learnt in this particular session, for example, was that participants were very concerned and one of their comments was, ''then we need to revisit our policies and procedures in the computerised environment''.

In the *Accounting* and *Human resources* sections, the focus was on transactions in the computerised environment vis-à-vis roles and responsibilities. We discussed at length the consequences of not having in place a detailed job description, in particular the issue of roles, responsibilities and accountability of staffs when dealing with various transactions such as financial in the computerised environment.

The effect of the awareness-raising sessions was a realisation of the need to go for further training. It triggered staffs, for example, to register for Certified Information Systems Auditor (CISA) and Certified Information Security Manager (CISM) after realising that they needed further training, even if it meant sponsoring themselves. CISA certification focuses on IT auditing, security, and control, while CISM focuses on the information security management (ISACA, 2005). It also triggered the concept of awareness through visiting other organisations (local and international) and preferably in the same industry, to learn what is going on as far as ICT security is concerned. Nevertheless, the local visits only showed that even the other organisations were still at the take-off stage.

2.8. *Step 8: carry out risk assessment and analysis*

Using the security team, we started to conduct risk assessment and analysis starting with the operations department (where core services of the organisation are located), followed by the IT department, physical security and later other departments. The cooperation from staffs was very high due to the effect of the awareness-raising sessions.

As suggested in Magnusson (1999), the need for countermeasures against ICT risks depends entirely on the effect these risks may have on the organisation's mission and business objectives. Fig. 5 is an extension of Fig. 1 and shows how these countermeasures are derived from the organisation's objectives.

(i) Identification of organisation's objectives
In Fig. 5, the objectives are represented by (O_1, O_2, O_3, O_4, …, O_n). The organisation's objectives, which will be taken into account, are those that are ICT dependent.

(ii) Identification of ICT assets that support the organisation's objectives
The second stage involves identification of ICT assets that support the organisation's objective/s (O_xA_x) and the business's key performance indicators. The ability of an organisation to achieve its mission and its business objectives is directly linked to the state of its ICT assets. As discussed in Alberts and Dorofee (2003), an asset is something of value to the enterprise and includes systems, information, software, hardware and people. Systems store, process, and transmit the critical information that drives organisations.

(iii) Analysis of threats to the organisation's ICT assets
The third stage involves threat analysis. For each identified asset, an assessment of the threats (A_xT_x) and their consequences that hinder the organisation from meeting its intended objective O_x takes place (where x identifies the objective and likewise the corresponding threat, and can be from 1 up to n threats). If we take the example of business continuity as an objective, then the set of threats can be theft, power fluctuation, virus or denial of service (DOS).

(iv) Ensuring organisation's objectives
The fourth stage involves identification of countermeasures for each threat. Picking theft in the previous example, the policy (P_x) may include backup, traceability and recovery, and user policy and procedures.

The outcome report (of identified objectives, identified ICT assets, threats and their possible countermeasures) is compared with the current organisation's ICT practices in order to estimate the security awareness in the organisation. The end result is the security benchmarking documented in a survey report that gives an overview of the security awareness and vulnerabilities in the organisation's ICT assets.

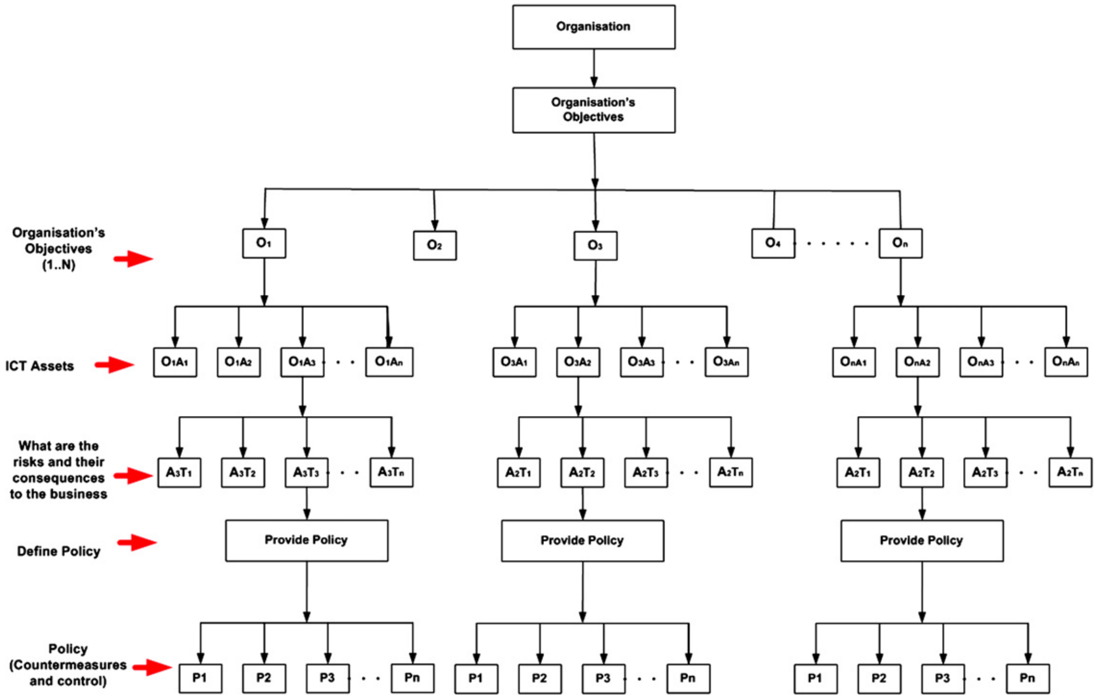


Fig. 5 – Showing how insurance policies can be derived from the organisation's objectives.

This exercise shed even more light on the magnitude of the security problem and information obtained from this step was vital for the discussion we held later with individual managers, in particular when discussing with the CFO on how to financially hedge the identified risks.

In addition, the obtained information was used to estimate security awareness when discussing with the IT department on which countermeasures are being practised and which are not. The discussion also took into consideration the output of the EMitL tool (the output of step 4) (Bakari et al, 2005b).

2.9. *Step 9: work out the mitigation plan (short-term plan for issues that need immediate attention and long-term mitigation plan)*

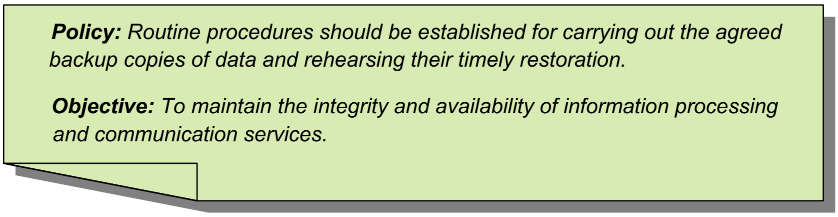
This is the step that came as a result of pressure from the management. Having realised how risky it was to go without proper ICT security management in place, the management was now at the forefront, suggesting that the security team come up with a mitigation plan. From the management side, an ICT steering committee (management focusing on ICT with security as a priority) was formed where management will work closely with the IT department. The need for control of information technology in use in the organisation was now realised as suggested in COBIT. We made use of OCTAVE method process 8 which involves developing a protection strategy to work out the mitigation plan. From the risk assessment and analysis and the quick scan that took place with the documentation, we found that there were issues that needed immediate attention. They included, for example, getting the issue of licences sorted out, patching the operating systems, training in some areas which were identified as critical but with not enough know-how, and improvement of the infrastructure which was also found to be part of the problem. Although all these were not budgeted for, the management saw the need to reallocate the budget for these immediately as they were seen to be cost effective, having a clear link in safeguarding the organisation's mission and business objectives.

A long-term plan was then worked out which included, among other things, a disaster recovery and business continuity plan, and developing countermeasures which included policies and various mechanisms including procedures on ICT security. These are detailed in step 10.

2.10. *Step 10: develop countermeasures*

The main question here was what set of countermeasures will provide the best protection against the identified risks and the state of ICT security in the organisation. The goal here is to design and develop countermeasures tailored to the organisation that will remedy the identified vulnerabilities and deficiencies. After this stage, which is mainly analytical, the solutions are still "on the drawing board", the process referred to in Information Security Management Systems (ISMS) (Bjorck, 2001). The operationalisation stage takes the conceptual level and makes it work in the organisation. This entails, for example, installation and configuration of technical security mechanisms (e.g. user policy and procedures, backup, etc.), as well as information security education and training of employees.

By taking into consideration the suggestion made from the EMitL tool (what should have been in place), ITIL (why), ISO 17799 (what), COBIT (how) and finally the environment in which the organisation is operating, we started deriving the



Policy: *Routine procedures should be established for carrying out the agreed backup copies of data and rehearsing their timely restoration.*

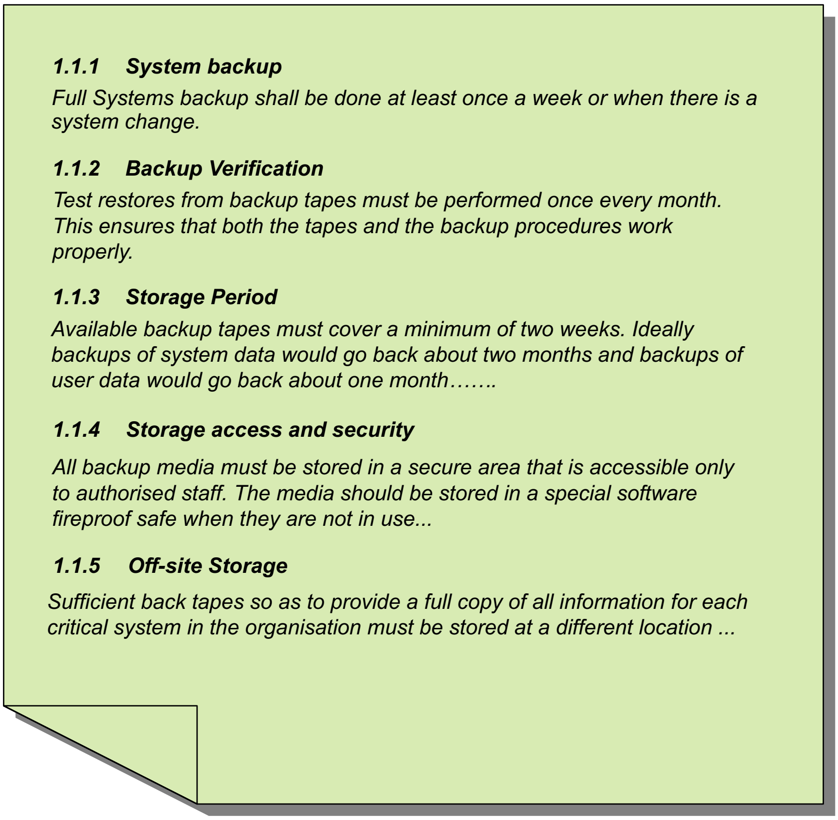
Objective: *To maintain the integrity and availability of information processing and communication services.*

Fig. 6 – Sample policy.

relevant countermeasures to be implemented in order to address the identified ICT risk exposure (ITIL, 2005; ISACA, 2005; ISO 17799). ITIL and COBIT helps in defining objectives of the processes involved in ICT organisation in general where ICT security is a part of the whole. We used ISO 17799 details on what should be done in addressing each dimension. For example, what should be done by human resource people to address the problem associated with ethics/culture. This is detailed under the subsection dealing with personnel security of the standard. Our approach was only to consider those issues that are basic and which can be achieved. This helped us to develop the ICT security policy and corresponding security mechanisms for the organisation. Fig. 6 shows part of the sample policy document where for each policy statement (what) there is the objective (why) which attempts to answer the question why the organisation should have such a policy statement.

For each policy statement we had the corresponding objectives and what type of vulnerability is being addressed. Then there is a pointer to the procedures which show in detail how such a policy is going to be implemented. For example, Fig. 7 shows *how* the above policy and objectives in Fig. 6 were translated into procedures.

This step triggered another concern of redefining job descriptions and responsibilities of the staffs in the organisation. The procedures and security mechanism we developed or suggested then became major inputs in this exercise. In addition, there was a reorganisation of the technical department to include the ICT security function. All these were driven internally through the ICT security project team (formed in step 3). This was achieved in two steps. First reorganisation of the IT department's organisational structure to ensure that there is a clear demarcation of responsibility. For example, system development was separated from the section that deals with change management, and the user department and the interface was centralised at the helpdesk. This was achieved by making use of ITIL (service management process) and ISO 17799 (personnel security). The second exercise



1.1.1 System backup

Full Systems backup shall be done at least once a week or when there is a system change.

1.1.2 Backup Verification

Test restores from backup tapes must be performed once every month. This ensures that both the tapes and the backup procedures work properly.

1.1.3 Storage Period

Available backup tapes must cover a minimum of two weeks. Ideally backups of system data would go back about two months and backups of user data would go back about one month…….

1.1.4 Storage access and security

All backup media must be stored in a secure area that is accessible only to authorised staff. The media should be stored in a special software fireproof safe when they are not in use...

1.1.5 Off-site Storage

Sufficient back tapes so as to provide a full copy of all information for each critical system in the organisation must be stored at a different location ...

Fig. 7 – Sample procedures.

involved brainstorming with the technical department on how the newly developed ICT security policy and procedures could be incorporated into the reviewed sections and the staff's roles and responsibilities. The activities identified in this second step were to wait until the following financial year. In addition, the plans for the new financial budget for each department took into consideration the operationalisation of the proposed countermeasures.

One issue that brought about some discussion was the positioning of the security function in the organisation. Our final conclusion for this, after discussing the advantage and disadvantage of positioning it in different departments, was to have ICT security positioned in the existing security department which was directly under the CEO's office and headed by the chief security officer with overall responsibility for ICT security. Another staff position that was created was that of ICT security administration at the IT directorate level.

Finally, we convened the management team to present the mitigation plan and the proposed countermeasures. One of the key messages that we delivered at the meeting was for them to take responsibility for ensuring that ICT security policy and procedures are approved by the board before full implementation starts. We brought to their attention that it is the responsibility of the board of directors and executive management to provide oversight of the implementation of information security (Posthumus and Solms, 2005), and therefore the outcome of this step (policy and procedures) should be brought to the attention of the board. It was the responsibility of the ICT security project team and IT steering committee to ensure that the policy and procedures come into operation.

3. Discussion

One of the major problems found in organisations today, including our case study, has to do with perception, where the management and general staffs perceive that ICT security is a technical problem and not a business problem. This situation leads to a misunderstanding of the nature of the security problem and consequently ends up in addressing the wrong problem. The observation has indicated that changing the way the management and general staffs perceive the problem is a necessary process and a prerequisite step towards a common understanding of managing ICT security in an organisation. This can be achieved through awareness-raising sessions. Furthermore, the backing and awareness of both management and the general staffs is vital for the success of the ICT security management programme, which leads to the appreciation that the protection of ICT assets is a business issue and not a technical issue.

In the case study, since a sense of security in non-digital (manual) processes exists in the management and staffs in general, what needs to be cultivated is a shift of focus from the manual to the digital process. For example, it is common, in particular in state-owned organisations, to have high security in place on how to handle confidential records in physical registries. This even includes a special type of recruitment of staffs (including vetting), who work in such offices. Being aware, for instance, that system administration is more sensitive than the mere registry, since system administrators have access to more information than that which is already printed, was another milestone. We also found that perception and interpretation of the words ICT security often leads to a misunderstanding of the actual ICT security problem. For example, the traditional interpretation of the word security for many people, in particular in Tanzania (where this study was conducted), meant something to do with physical security, the police, etc., and the word ICT refers to modern technology only. The two key words *ICT* and *Security* should therefore be used carefully. When discussing ICT security with the management, it may sound better if we used Managing technology risks instead of Managing ICT security. Similar suggestions are found in Blakley et al. (2001) where information security is viewed as information risk management. Our experience is that staffs tend to be more cooperative and proactive in dealing with the problem when they understand exactly what ICT- related risks are all about. For instance, sharing passwords or issuing passwords to fellow staff members was not considered such a big deal, because not many staffs realised how dangerous that was.

ICT security is a multidimensional discipline consisting of, among other things, legal, human resources, technical, operations, security, audit, insurance, and finance (Solms and Solms, 2004). It is therefore important that the initialisation process, which involves the formation of a special project team, starts with the right staff. As discussed in the paper, for better results the team must be made up of senior staffs from all major departments (legal, human resources, technical, operations, security, audit, insurance, and finance) to be able to meet the multidimensional requirements of the ICT security problem. Another important aspect is to have practical examples during the awareness-raising sessions coming from the organisation itself, when discussing the ICT-related risks with the management and general staffs. This also helps when discussing the information security plan which must be based on the risk exposure of the organisation itself.

Getting the current status of ICT security of the organisation documented properly gives more knowledge, not only to the security team, but also to the management and general staffs of the interrelationship of ICT assets, threats, and vulnerabilities, as well as the possible impact on the organisation's mission and business objectives. It helps the management and general staffs appreciate the ICT security problem and hence assist in making them more supportive when dealing with the problem. In addition, awareness is very essential to all users but, as discussed before, it will have a significant impact if it is extended and the matter brought to the specific attention of different departments in the organisation. For instance, when meeting corporate lawyers, the main discussion will be on ICT-related risks from a legal point of view.

There are many internationally established codes of practice that are essential in the process of managing information security in an organisation. Studying these multiple sets of practices and guidelines is of importance for determining and understanding the features that are being recommended to organisations and which must be considered when managing information security. In our study, an attempt to approach the problem holistically was used, by initially merging two holistic approaches, BRITS and SBC model. BRITS gives the business view of the problem and the SBC the security by consensus. The result of the merger was used to approach the problem from the management's perspective as shown in step 5. We have used OCTAVE, ITIL, ISO 17799 and to some extent COBIT, in an attempt to compensate for the missing links in different steps. Looking at these three approaches, ITIL addresses ICT services and operational management practices that contribute to security, COBIT addresses control objectives for information technology security and process control and ISO 17799 is exclusive to the information security management process. A similar study by Solms (2005) has already shown the synergy of combining more than one framework when attempting to manage ICT security in an organisation.

Reviewing the steps as described here, it becomes apparent that they fit well with the issues discussed, such as a framework for information security governance and the like, and those discussed in the 10 deadly sins of information security management (Solms and Solms, 2004; Posthumus and Solms, 2004), although the order in which they appear here might be different.

The process (10 steps) needs to be initiated from outside, but then there is a need to have the process driven internally. Not many organisations have the capability of putting together the ingredients from different methods. An expert is required to interpret and apply different methods and apply what is required in specific stages. Our experience, however, indicated that it is possible to address this problem by extending the proposed holistic approaches into a set of guidelines which can be used to address the problem in the organisation.

4. Conclusion and reflections

Our objective to bridge the gap between the management and the technical department was achieved through the 10 steps. These included: the CEO buying into the idea first; recognising that the technical departments are the custodians of ICT in the organisation; starting it as a special project; showing where the risks and their consequences are; getting the entire management's attention; taking stock of the existing situation; conducting awareness-raising sessions to address the ICT security problem with respect to the organisation's specific environment; carrying out detailed risk assessment; working out a short-term plan for issues that need immediate attention and a long-term plan to finally develop the countermeasures for the identified problems. The study confirmed that the success of the ICT security management process begins with the management realising the importance of ICT security management. That implies that the management allows the organisation, through its own acquired knowledge and confidence, to internalise the practices, thus enabling people to act confidently at all levels. Knowing about the ICT risks and their consequences for the core service operations of the organisation, the management is more likely to offer its support for ICT security endeavours. Likewise, the technical department, following the support from the management, can address the ICT security problem more holistically in collaboration with other departments, by taking into consideration the non-technical dimensions as well.

Discussing bridging of the gap between the management and the technical department in general would also involve other stakeholders as well as looking at other angles of the problem. Within the Tanzanian experience, part of the research into ICT security has covered, in particular, ICT security education and training, ICT security management, Security Controls implementation and ICT systems security assurance (Casmir, 2005; Bakari, 2005; Tarimo, 2003; Chaula, 2003). These are all ongoing activities that hopefully will enable the country to find useful, efficient and socially acceptable ways of balancing the two main perspectives; the social (cultural and structural) and the technical (machines and methods) towards controllable, viable and homeostatic states.

REFERENCES

Alberts C, Dorofee A. Managing information security risks: the OCTAVE approach. Addison Wesley, ISBN 0-321-11886-3; 2003.

Bakari JK. Towards a holistic approach for managing ICT security in developing countries: a case study of Tanzania. Ph.L. thesis, SU-KTH, Stockholm. DSV report Series 05-011; 2005.

Bakari JK, Tarimo CN, Yngström L, Magnusson C. State of ICT security management in the institutions of higher learning in developing countries: Tanzania case study. In: The 5th IEEE ICALT, Kaohsiung, Taiwan; 2005a. p. 1007–11.

Bakari JK, Magnusson C, Tarimo CN, Yngström, L. Ensuring ICT risks using EMitL tool: an empirical study, IFIP TC-11 WG 11.1 & WG 11.5 joint working conference on security management, integrity, and internal control in information systems, December 1–2, Fairfax, Virginia, Washington, US; 2005b. p. 157–73.

Bishop M. Computer security, art and science. Addison Wesley, ISBN 0-201-44099-7; 2003.

Bjorck F. Security Scandinavian style, interpreting the practice of managing information security in organisations. Ph.L. theses, Department of Computer and Systems Science, University of Stockholm and the Royal Institute of Technology, Stockholm; 2001.

Blakley B, McDermott E, Geer D. Information security is information risk management. In: Proceedings of the 2001 workshop on new security paradigms. New York, NY, USA: ACM Press; September 2001.

Casmir R. A dynamic and adaptive information security awareness (DAISA) approach. Ph.D Thesis, SU-KTH, Stockholm; 2005. No. 05-020.

Chaula JA. Security metrics and public key infrastructure interoperability testing. Ph.L Thesis, SU-KTH, Stockholm, DSV report Series 03-021; 2003.

Caralli AR. Managing for enterprise security. USA: Carnegie Mellon University; December 2004.

ISACA. <http://www.isaca.org/cobit/>; 2005 [last accessed on 20 October 2005].

ISO 17799 Standard.

ITIL. <http://www.itil.org.uk/>; 2005 [last accessed on April 2005].

Kowalski S. IT insecurity: a multi-disciplinary inquiry. Ph.D. Thesis, Department of Computer and Systems Sciences, University of Stockholm and the Royal Institute of Technology, Stockholm; 1994. ISBN: 91-7153-207-2.

Magnusson C. Hedging shareholders value in an IT dependent business society. The framework Brits. Ph.D Thesis, Department of Computer and Systems Science, University of Stockholm and the Royal Institute of Technology, Stockholm; 1999.

Solms BV, Solms RV. The 10 deadly sins of information security management. Computers & Security 2004;23(5). ISSN: 0167-4048:371–6.

Solms BV. Information security governance: COBIT or ISO 17799 or both? Computer & Security 2005;24:99–104.

Posthumus S, Solms RV. A framework for the governance of information security (Elsevier Ltd.). Computers & Security 2004;23:638–46.

Posthumus S, Solms RV. A responsibility framework for information security. In: IFIP TC-11 WG 11.1 & WG 11.5 joint working conference on security management, integrity, and internal control in information systems, Fairfax, Virginia, Washington, US; 1–2 December 2005. p 205–21.

Tarimo C.N. Towards a generic framework for implementation and use of intrusion detection systems. Stockholm University/Royal Institute of Technology, Report series No. 2003-022, SU-KTH/DSV/R – 2003–SE; December 2003.

Wilson M, Hash J. Building an information technology security awareness and training program. NIST Special publication 800-50; October 2003.

Jabiri Kuwe Bakari is a Ph.D. student studying potential solutions in relation to the management of ICT security (holistic approach), at the Department of Computer and System Sciences, Stockholm University/Royal Institute of Technology in Sweden. He received his B.Sc. Computer Science degree at the University of Dar-es-Salaam Tanzania in 1996, M.Sc. (Eng.) Data Communication degree from the Department of Electronic and Electrical Engineering, Sheffield University in UK in 1999, and Licentiate of Philosophy degree from the Department of Computer and System Sciences, Stockholm University/Royal Institute of Technology in Sweden in 2005. He is an active member of the International Federation for Information Processing (IFIP) TC-11 Working Group 11.1, and IEEE. He has published and presented several papers in the field of information security management at the ISSA, IFIP, IEEE and IST international conferences.

Charles Tarimo is currently a doctoral candidate in computer and communication security at Stockholm University, Department of Computer and Systems Sciences. He holds a B.Sc. in Engineering (B.Sc Eng.) obtained in 1994 from the University of Dar-es-Salaam, Tanzania and a Licentiate of Philosophy (Ph. lic.) in Computer and Systems Sciences, obtained in 2003 from Stockholm University in Sweden. Charles is an employee of the University of Dar-es-Salaam Tanzania. His research interests are focused on operational and practical issues with regard to aspects of requirement development, designing, implementation, and maintenance of different technical and non-technical ICT security controls within organisations, such as Intrusion Detection Systems.

Louise Yngström is a Professor at the Department of Computer and Systems Sciences, Stockholm University/Royal Institute of Technology. She is also the Director of SecLab, Dean of research studies, and responsible for national and international masters programmes in ICT security. She started one of the very first interdisciplinary academic IT security programmes in the world in 1985, naming it ''Security Informatics''. She was awarded her Ph.D. in 1996 for the introduction of a methodology for such academic programmes, called the ''Systemic-Holistic Approach'' (SHA), where ''soft'' and ''hard'' sciences appropriate for IT security issues are mixed. Being one of the pioneers in the field of systems sciences and security in Sweden, she has been with the department since 1968. Dr. Yngström founded IFIP's WG11.8 and the World Conference on Information Security and Education, and is an active member of various working groups within IFIP TC9 (Social accountability of IC&T) and TC-11 (Computer Security).

Christer Magnusson is an Assistant Professor at the Department of Computer and Systems Sciences, Stockholm University/Royal Institute of Technology, specialising in IS/IT Security and IS/IT Risk Management. He brings two decades of industrial and academic information security experience to our group. Before joining SecLab, Dr. Magnusson was the Head of Corporate Security and Risk Management at Sweden Post and CEO of Sweden Post Insurance AB, and he has also been the Head of Corporate Security in the Ericsson group. He has also worked within the IT Security group of the Swedish Agency for Public Management (Statskontoret). Dr. Magnusson was awarded the SIG Security Award by the Swedish Computer Society in 1999 and in 2000 the Security Award by the Confederation of Swedish Enterprise (Svenskt Näringsliv) in recognition of the models and the integrated processes regarding IS/IT Risk Management that he developed as a part of his research studies. He holds M.Sc. and Ph.D. degrees in Computer and Systems Sciences.

Dr. Stewart Kowalski is a part-time lecturer and advisor at the Department of Computer and System Sciences, Stockholm University/Royal Institute of Technology in Sweden. He has over 25 years experience in teaching and IS/IT security. He is currently the Risk Manager for Ericsson Global Services which operates in over 140 countries around the world. His research interests include industrial and academic research in the development adoption and assimilations of IT security technologies and practices in organisations, markets and cultures.

Computers
&
Security

Organisational security culture: Extending the end-user perspective

A.B. Ruighaver[*], *S.B. Maynard, S. Chang*

Department of Information Systems, University of Melbourne, Australia

Keywords:
Security culture
Organisational culture
Security governance
Motivation
Risk analysis

ABSTRACT

The concept of security culture is relatively new. It is often investigated in a simplistic manner focusing on end-users and on the technical aspects of security. Security, however, is a management problem and as a result, the investigation of security culture should also have a management focus. This paper describes a framework of eight dimensions of culture. Each dimension is discussed in terms of how they relate specifically to security culture based on a number of previously published case studies. We believe that use of this framework in security culture research will reduce the inherent biases of researchers who tend to focus on only technical aspects of culture from an end-users perspective.

1. Introduction

It was not until the start of this century that researchers first began to recognise that an organisation's security culture might be an important factor in maintaining an adequate level of information systems security in that organisation (Schwarzwalder, 1999; Breidenbach, 2000; Von Solms, 2000; Andress and Fonseca, 2000; Beynon, 2001). None of these authors, however, presented a clear definition of what they meant with ''a security culture'', nor were there any clear views on how to create this organisational culture to support security.

In the last few years, research in this new area of (information) security culture has been expanding rapidly. Unfortunately, a lot of this research still has a limited focus and often only concentrates on the attitudes and behaviour of end-users as well as on how management can influence these aspects of security culture to improve the end-user's adherence to security policies (Schlienger and Teufel, 2002; Ngo et al., 2005). Schlienger and Teufel (2003) more or less defines security culture as ''all socio-cultural measures that support technical security measures'', which not only limits its focus to a small sub-dimension of information security, but also enforces the old belief that information security is mostly a technical problem. Information security is, in general, a management problem and the security culture reflects how management handles this problem. Subsequently, we will argue that technical security measures and security policies will often need to be (re)designed to support an organisation's security culture.

In this paper we propose that security policy development is just one of the areas that will be influenced by an organisation's culture. Nosworthy (2000), for instance, states that an organisation's culture has a strong influence on organisational security, as it may 'hinder change'. Borck (2000) states that 'beyond deploying the latest technology, effective security must involve the corporate culture as well'. There is strong suggestion from the literature that the study of security culture cannot be carried out in isolation of wider organisational culture. For example, in their review of organisational

[*] *Corresponding author.*
E-mail addresses: anthonie@unimelb.edu.au (A.B. Ruighaver), seanbm@unimelb.edu.au (S.B. Maynard), shanton.chang@unimelb.edu.au (S. Chang).

behavioural studies, Mowday and Sutton (1993) point out that contextual factors play a significant role in influencing individual and group behaviours within organisations. The contextual factors here often reflect the organisation's culture.

In the following sections we will describe how we used Detert et al.'s (2000) framework of eight 'overarching, descriptive culture dimensions' to explore the security culture within quite a few organisations with vastly different levels of security. As several of these case studies have been published previously (Shedden et al., 2006; Maynard and Ruighaver, 2006; Koh et al., 2005; Tan et al., 2003; Chia et al., 2003, 2002), we will concentrate on the resulting insights that these case studies have given us into each of these dimensions of an organisational security culture.

2. Exploring organisational security culture

Our initial research in organisational security culture adopted a framework with eight dimensions from Detert et al. (2000). Detert et al. (2000) synthesised the general dimensions of organisational culture using current organisational culture research on areas such as Organisational Culture and Leadership (Schein, 1992), Competing Values (Cameron and Freeman, 1991) and Organisational Culture Profile (Klein et al., 1995). Detert et al. (2000) illustrate their framework by linking it to a set of values and beliefs that represent the 'cultural backbone' of successful Total Quality Management (TQM) adoption. For a group of security experts with limited initial knowledge of organisational culture, this clear description of the cultural aspects of TQM convinced us of the power of this framework in exploring security culture. The eight dimensions of organisational culture are briefly identified in Table 1.

3. Interpreting organisational security culture

In the remainder of this paper we give our current views of what the important aspects are of security culture for each of these dimensions. These views have been constructed over a number of years and have been influenced by the case studies the authors have completed in various aspects of security including governance and security culture (Shedden et al., 2006; Maynard and Ruighaver, 2006; Koh et al., 2005; Tan et al., 2003; Chia et al., 2003, 2002). In some of the case studies, the security culture was observed as part of understanding other aspects of security, whilst in others, there was a specific focus on the security culture of the organisation. While a few of our case studies have been in organisations that have a high level of security enforced by a strict enforcement of rules and regulations, the majority of our research has been in organisations where decision making about security is distributed and loosely controlled. Whilst this may have slightly coloured our views expressed below, the inclusion of organisations with strict security as well as those with less strict security allows this research to be informed by different organisation types and thus different types of organisation and security culture.

The rest of this section uses each of the dimensions of the security culture model to describe the important aspects of security culture.

3.1. *The basis of truth and rationality*

What we initially considered to be our most important findings in our early research on security culture related to how the importance of security for the organisation is seen by the employees and the organisation as a whole. Obviously, different organisations need different levels of security, but although the security requirements for a particular company may not be as high as the security requirements of other companies, achieving optimal security for that organisation's particular situation will still be important; as is the need to ensure that their employees believe that security is important. While the literature on security culture recognizes that the most crucial belief influencing the security in the organisation is the belief, by both employees and by the organisation itself, that security is important (Connolly, 2000), not much is mentioned about the importance of the other beliefs that an organisation may have about security.

After more extensive research (Shedden et al., 2006), we found that any beliefs that the decision makers within the organisation have about the quality of security, and about the quality of the different processes used to manage security, are often much more important than the end-user's beliefs about security. Many of the organisations investigated, whether they have low or high security requirements, believe that their security is good. However, most of these organisations do not really make any attempt to evaluate the quality of their security, or even attempt to measure its success, except anecdotally. Similar problems exist with the organisations' beliefs about the quality of their risk analysis and security audits.

In those case study organisations where there is a high degree of security, the assessment of the quality of security tends to focus on the trust of the extensive security processes in terms of policy and procedures. However, in those organisations with less emphasis on security it is clear that their assessment that the security within the organisation is good, may often be flawed. Even though they have a lower requirement for security these organisations tend to treat security spending as a cost to the organisation, and it often is a struggle to gain funding for security initiatives as a result. This tends to send out conflicting messages to employees. On one hand, management states that security is important to the organisation, however, on the other hand, it is not willing to support and fund security initiatives. Thus a conflict may occur as to what is the truth in the organisation regarding security.

The quality of a security culture should, however, not only be determined by the beliefs that an organisation has, but more by how the organisation evaluates and manages the basis of truth and rationality in the various beliefs that end-users and managers hold about that organisation's security. Staff being critical about their own beliefs and an organisation having processes in place to challenge the quality of the beliefs of its employees is what distinguishes a good security culture from a bad one.

Table 1 – The organisational culture framework (Detert et al., 2000)

1. *The basis of truth and rationality*
   Focuses on the degree to which employees believe something is real or not real and on how the truth is discovered. This dimension may affect the degree to which people adopt either normative or pragmatic ideals.

2. *The nature of time and time horizon*
   The concept of time in an organisation has baring in terms of whether the organisation adopt long-term planning, strategic planning and goal setting, or focus primarily on the here and now, reacting on a short time horizon.

3. *Motivation*
   Within an organisation motivation is a fundamental management principle. The identification of how employees are motivated; whether they are motivated from within or by external forces is important. Furthermore, whether employees are inherently good or bad, whether they should be rewarded or punished, and whether manipulating others' motivation can change effort or output are all characteristics of motivation.

4. *Stability versus change/innovation/personal growth*
   Stability and change are closely linked to motivation. Some individuals are open to change (risk-takers), whereas other individuals have a high need for stability (risk-averse). This can also apply to organisations. Risk-taking organisations are said to be innovative with a push for constant, continuous improvement. Risk-averse organisations tend to be less innovative, with little push for change.

5. *Orientation to work, task, co-workers*
   The centrality of work in human life and the balance between work as a production activity and as a social activity. Some individuals view work as an end in itself and are, concerned with work accomplishment and productivity. Other individuals see work as a means to other ends, such as having a comfortable life and developing social relationships. Issues such as the responsibility employees feel for their position and how they are educated in terms of their roles and responsibilities are important here.

6. *Isolation versus collaboration/cooperation*
   Focuses on how employees can work, either alone, or collaboratively. Underlying beliefs about the nature of human relationships and about how work is most effectively and efficiently accomplished. In some organisations the majority of work is accomplished by individuals, and collaboration is often viewed as a violation of autonomy. Other organisations welcome collaboration and foster team work, often organizing work around groups of employees.

7. *Control, coordination and responsibility*
   Organisations vary in the degree to which control is concentrated or shared. Where there is firm control there are formalized rules and procedures that are set by a few, to guide the behaviour of the majority. Where there is less control there is flexibility and autonomy of workers, with fewer rules or formal procedures and shared decision making.

8. *Orientation and focus – internal and/or external*
   The nature of the relationship between an organisation and its environment and whether or not an organisation assumes that it controls, or is controlled by, its external environment. An organisation may have an internal orientation (focusing on people and processes within the organisation) or external orientation (focusing on external constituents, customers, competitors and the environment), or have a combination of both.

3.2. *Nature of time and time horizon*

As literature already indicated (Wood, 2000), we found that all too often the security focus of an organisation is on things demanding immediate attention, not on the things that may prove more important in the long run. In many of the organisations we investigated, those that had any long-term goals tended to only cover a timeframe of one or two years and were primarily aimed at building a solid security infrastructure in line with International Security Standards. One common theme in those organisations focusing on the short term was expressed by one of the respondents (a manager level employee) who stated that "unless there is a breach by someone, security is never addressed again after the initial training". This is in contrast to an organisation with a longer term strategy, even where that timeframe is one to two years, which consistently reminds employees of their responsibilities to security through various organisational processes such as, mail out reminders, notice boards and security bulletins.

Although we found no good examples in our case studies, we still argue that organisations with a high-quality security culture should place an emphasis on long-term commitment and strategic management. Unfortunately, there is not much discussion in literature on possible long-term strategies either. There seems to be a tendency, however, to completely overhaul security management/governance structures when current security is no longer adequate and/or becomes too expensive to maintain. When that happens, once again there does not seem to be any evidence that those initiating this restructuring have even considered what long-term strategies and plans can or should be developed and by whom. When the authors suggest an obvious (in their view) long-term strategy aimed at building up appropriate skill-sets related to security, or aimed at aligning the organisation's security with its organisational culture, these suggestions are often thrown out as being too expensive and not requiring a high priority.

3.3. *Motivation*

From the case studies we found no evidence that employees are intrinsically motivated to adopt secure practices. Hence, organisations with a good security culture need to have appropriate processes in place to ensure employees are motivated in relation to security. While literature suggests that employees need to learn that security controls are necessary and useful to discourage them from attempting to bypass these controls (Lau, 1998), motivation should not just be aimed at ensuring that an employees behaviour is not compromising IS security.

Additionally, in looking at the impact of employee motivation on eventual behaviours, literature has maintained predominantly that provision of extrinsic rewards to employees for performing particular tasks may actually damage their intrinsic motivation for performing the tasks (Eisenberger and Cameron, 1996). However, more recent research has indicated that tangible rewards may be linked to positive intrinsic motivation amongst employees, particularly where there are changes and increases expected in performance standards (Pierce et al., 2003).

In examining security culture, end-users would be expected to meet slightly modified performance standards and reconsider their accepted behaviours. In some instances, they may be asked to modify these behaviours. Hence, there is a suggestion from the literature that it may be important to consider both tangible rewards to positive behavioural change (e.g. money) and intrinsic motivation to adopt new behaviours (e.g. recognition and social participation). Furthermore, it will be important for organisations not to contradict the values that they are trying to impose on employees with regard to security culture by setting managerial goals that short circuit security goals. For instance, it may be the case that management key performance indicators are being set to encourage and reward behaviour, but contradict the goals the organisation is trying to achieve with security.

In an ideal security culture, end-users, security administrators and managers will be motivated to reflect on their behaviour at all times, to assess how their behaviour influences security and what they can do to improve security. To create this attitude to security, it is important that a degree of trust is involved and that responsibility to act in an appropriate manner is delegated to employees themselves. However, this does not mean that an organisation should not have monitoring processes in place to identify security breaches, to ensure that unacceptable behaviour is corrected, and to reward exemplary behaviour. Furthermore, motivation may occur through employees having ownership of security. In some case study organisations, the responsibility of security was passed from a management level through to each and every employee in the organisation. As such, peer group pressure became a driving force in one of the organisations for employees to abide by security procedures and employees were motivated to be secure in their work practices.

As a result of a recent case study (Koh et al., 2005), we believe that the best way to improve motivation may be through horizontal social participation. While it is common in organisations to encourage social participation as in encouraging staff influenced by a decision to participate in the decision making process (which we call vertical social participation), such limited social participation has only a limited effect on improving the security culture. Conversely, there are numerous instances within organisations where people at the same level within different areas come across the same security issues and may not know that others in the organisation are covering the same ground. Organisations that have a more broader form of social participation in which, for instance, all system and security administrators across the business units are involved in exchange of information to improve decision making, may find that motivation (as well as many other cultural aspects) will increase significantly.

3.4. *Stability versus change/innovation/personal growth*

An important aspect of examining an organisation's culture is to look at its tolerance for change and innovation. From the case studies it is clear that in some organisations, periodic cycles of change are purposefully built into their culture and processes (Brown and Eisenhardt, 1998). In such organisations, new products and services are introduced which require new processes and flexible structures within the organisation. Consequently, such organisations have a culture where individual risk taking behaviour within acceptable boundaries may be tolerated or even encouraged (Brown and Eisenhardt, 1998).

In many of the case study organisations that have lower security requirements, they have a propensity to be reactive to change rather than proactive. For instance many organisations of this type in the case studies tended not to have any security procedures and policies for dealing with privacy issues until legislation was passed by the Australian government forcing them to deal with this issue. The more security conscious organisations studied tended to be more proactive and whilst they still needed to react to this issue, they tended to have procedures already in place.

Whilst organisations that have lower requirements for security often are tolerant to change, they often fail to realize that an organisation's security procedures and practices need to improve continually, and that the organisation will need to constantly adapt its security to the inevitable changes in the organisation's environment. Furthermore, many of these types of organisations studied tended to overlook the importance of change management in terms of security. Their tolerance for change tended to affect the security of the organisation as changes were often implemented without sufficient study as to their impact on security.

In those case study organisations that have a high requirement for security, we found a tendency to favour stability over change. Change is often seen as bad, as it can result in the introduction of new risks or in the invalidation or bypass of controls to existing risks. In this type of organisation change is often a formal process and is well managed and well controlled. In urgent situations, the process is often fast tracked through meetings, rather than through the completion of the formal change process. Although change should be carefully managed, security is never 100% and organisations therefore need to ensure that their 'security posture is not static' (Shinn, 2000). ''Security change management is an ongoing process, daily, you couldn't put a time to it, it takes place all the time, it's constant'' (a CIO case study participant).

Those organisations that have adopted a security policy lifecycle methodology will have a culture of continuous change in that area of security, but it is not clear whether this will extend to other areas such as security strategy development and security governance processes, or even implementation of security measures.

An aspect of security culture that we found lacking in almost all of the case study organisations is the development of new and innovative approaches to security. Most organisation just use the same old traditional security technologies and controls, not taking advantage of the fact that every organisation is different and that new challenges in their security environment may warrant new and unconventional approaches.

3.5. *Orientation to work, task, co-workers*

An important principle in information security is that there is always a trade-off between the use of an organisation's assets and their security. By limiting access to an asset, we can significantly improve its security. However, limiting access can sometimes result in a serious impediment to the daily operations of employees. In all organisations studied it is clear that a trade-off occurs, but it is interesting to note that in those organisations that have high security requirements, employees tend to more readily accept the limitations placed on the organisation's assets because of security concerns. In some of the organisations with lower security requirements, employees became resentful of the security restrictions imposed, even though they are much less than in a highly secure environment. Finding a balance between security and how constrained employees feel in their work is therefore an important aspect of a security culture. Of course, staff will feel less restricted if they are motivated and feel responsible for security.

While it is obvious that employees should be made to feel responsible for security in the organisation, it is just as important that staff responsible for particular security areas have as strong sense of ownership. This will again be influenced by the amount of social participation staff has in security (Koh et al., 2005), but can be negated easily when staff feel that management do not take any suggestions for the improvement of security very seriously. Hence, a positive response of management and a continuous adaptation of security practices to at least some of the suggestions may not only help improve security itself directly but also help improve the orientation of staff towards security.

Education of employees on their roles and responsibilities related to security is also crucial (Freeman, 2000). Too many organisations only give employees an overview of security during induction, and even then they mostly cover aspects of what is considered acceptable use. Adequate user education can 'eliminate inadvertent disclosures, and even help users contribute to total system security' (Hartley, 1998). Furthermore, education is also an important tool in increasing the feeling of responsibility and ownership of those staff involved in decisions about security and involved in implementing those decisions. But for such education to have a significant impact on the employees' orientation to work it needs to be reinforced continuously, and needs to respond to any unsatisfactory behaviour that has become widespread enough for users to consider it normal behaviour.

3.6. *Isolation versus collaboration/cooperation*

We have been surprised after conducting the case studies how often we encountered that an organisation's security planning and implementation was handled by only a small group of specialists and managers. In all organisations it was clear, however, that having a single person with these responsibilities was not acceptable. While organisations often realize that security policies should be created collaboratively using the input of people from various facets of the organisation to ensure its comprehensiveness and acceptance (Clark-Dickson, 2001), they tend to ignore that principle in the day to day management of security as well as in the design of organisational security processes. As a result, the efforts of the security management team are often negated by other decisions taken by managers in the business units and on the work floor. For instance, in one case study organisation, the security team mandated that usernames and passwords for systems would not be shared. However, in the day to day operations of the organisation, the processes involved in maintaining customer records meant that often passwords to access particular customer accounts were shared between account agents, with knowledge and express permission of their line manager to do so.

Our current research in security governance processes and structures at the middle management level is indicating that this lack of collaboration with the stakeholders in the day to day decision making on security is not only likely to negatively impact motivation and orientation to work, but may often also lead to a dangerously narrow focus of security. As coverage is just as important in information security as the quality of the selected security controls, ignoring particular areas such as personnel security or data security can lead to a significant collapse of an organisation's security posture.

3.7. *Control, coordination and responsibility*

This dimension of an organisation's security culture is clearly related to the organisation's security governance and has been the main reason that our security group extended its research from security culture to security governance. As discussed previously, security governance at the middle management level is often limited to a small group of people leading to isolation of the decision makers. In this section we will mainly discuss the implications of security culture on the way decisions on security are made.

The primary feature of security governance in an organisation is whether there is a tight control or loose control. An organisation with centralised decision making tends to have a tight control, while an organisation that has flexible decentralised decision making is likely to have a loose control, although change management processes may still influence how loose the control actually is. It should be clear that security culture is not independent from organisational culture, so tight control of security in an otherwise loosely controlled organisation is not likely to work very well.

The case study organisations range from extremely tightly controlled organisations through to those where there is considerable leeway in control. It is evident in the case study organisations that where an organisation has high security requirements they are also likely to have tight control over processes and policies. Likewise many of the organisations with lower security requirements tend to have less control over their procedures and policies. Furthermore, in organisations where there is a misalignment of security goals and organisational goals there is likely to be reluctance in supporting security initiatives, and as a result the organisation may have little co-ordination of security.

Whether there is tight or loose control over the decision making process of the organisation will also influence the level to which an organisation tolerates individual initiative. In some ways, this reflects the level of formalization within an organisation. Baker and Jennings (1999) indicated that even where there are mechanisms of control and formalization within the organisation, a culture of fear, uncertainty and loose control may be such that these control mechanisms such as policies, rules and procedures can be rendered dysfunctional.

However, independent of whether there is a tight control or a loose control, it is still essential that there are clear guidelines on who has decision rights in the different areas of security and when. This aspect is often called responsibility, and ensuring that all responsibilities have been assigned is a required feature in any strategic security policy. It should be realised, however, that having responsibility and feeling responsible are two different issues.

With responsibility comes accountability. We believe that an important aspect of the security culture is how the organisation handles accountability for decisions in security management. Lack of even the most simple accountability processes, such as simple feedback loops where decisions are discussed with higher levels of management, is a fairly common occurrence in security management.

In one case study organisation (Koh et al., 2005), all participants were asked about the areas of security they wished to improve. Interestingly there was a one to one correlation between these areas of security improvement with each individual's responsibilities to security within the organisation. This may indicate that there is a high level of accountability and responsibility for security, especially when responsibilities (and authority) are delegated throughout an organisation.

Additionally, the delegation of responsibility to employees does not preclude the need for top management support. Knapp et al. (2006) found that top management support for information security is a significant predictor of both the direction of an organisation's security culture and the level to which its security policies are enforced. Therefore, whereas operational responsibility and accountability lies with middle management and end-users, top management has a clear responsibility to:

- visibly demonstrate a prioritization of information security,
- take security issues into account in planning organisational strategies, and
- provide strong and consistent support to the overall security program.

3.8. *Orientation and focus – internal and/or external*

The orientation and focus of an organisation's security clearly depends on the environment in which the organisation operates. From the case studies, there are a number of organisations that are forced to conform to external audit and government requirements. As such, the emphasis of their risk management processes is often only on meeting these requirements, not on improving their security (Maynard and Ruighaver, 2006). Other case study organisations aimed to bring their IS security in line with international industry standards. Once again, the emphasis was geared towards passing an audit to prove that they have achieved this goal, rather than on achieving the best security for the organisation within the obvious limitations of resources and budget. In each of these cases the focus of the organisation was on external forces, and they often neglected to have an internal focus.

Some of the case study organisations, those that tended to have medium to high levels of security, had a focus that was both inward and outward looking. In these types of organisations, whilst it was important to conform to legislative requirements, and often to standards, they undertake these tasks whilst looking internally to determine how best the organisation can benefit from these things. As a result, organisational processes, policies and procedures are positively influenced.

As security in an organisation is influenced by both external factors and internal needs, we believe that an ideal security culture has a balance between an internal and external focus, much like the aforementioned organisations. The external focus should at least include an awareness of the organisation's external security environment and how this changes over time. This will allow the organisation to pro-actively meet any new threats. More important, however, is that the organisation builds up an awareness of its internal security environment. If the organisation is not trying to identify what security breaches occur and why they occur, it will never know if its security strategies are working and how it can improve the implementation of these strategies.

4. Conclusion

While there has been an abundance of research in the area of organisational security and how it should be improved, most only focus on certain aspects of security and not how these aspects should be assimilated (or integrated or taken into account) into an organisation's culture. Even our own research in security culture initially had a clear bias to end-user issues. However, the broad culture framework we adopted from organisational culture research has ensured that we not only recognised this bias in our research, but also has provided insight in how to extend our security culture research in new areas such as security governance and risk assessment.

In investigating security cultures in organisations, we have often found that many specific aspects of a security culture, such as attitudes, norms, shared expectations and many more, in general do not fit nicely within a single dimension of our framework. It is obvious that the concept of a security culture is too complex to be covered by a single framework or model. While we are, therefore, still hesitant to give a definite opinion of what a good security culture is, or even to give a definition of what exactly the concept of a security culture is, we do believe that any researcher involved in investigating any aspect of an organisation's security culture will find the use of Detert et al.'s (2000) framework essential in ensuring that they take a comprehensive view of how the different dimensions of an organisation's security culture relate to that particular aspect they are interested in.

REFERENCES

Andress M, Fonseca B. Manage people to protect data. InfoWorld 2000;22(46):48.

Baker EH, Jennings K. Dysfunctional organisational control mechanisms: an example. Journal of Applied Management Studies 1999;8:231–3.

Beynon D. Talking heads. Computerworld 2001;24(33):19–21.

Borck J. Advice for a secure enterprise: implement the basics and see that everyone uses them. InfoWorld 2000;22(46):90.

Breidenbach S. How secure are you? Information Week 2000;800: 71–8.

Brown SL, Eisenhardt KM. Competing on the edge: strategy as structured chaos. Boston, Massachusetts: Harvard Business School Press; 1998.

Cameron K, Freeman S. Cultural congruence, strength and type: relationships to effectiveness. Research in Organisational Change and Development 1991;5:23–58.

Chia P, Maynard S, Ruighaver AB. Understanding organisational security culture. In: Sixth pacific Asia conference on information systems, Tokyo, Japan; 2–3 September 2002.

Chia P, Maynard S, Ruighaver AB. Understanding organisational security culture. In: Hunter MG, Dhanda KK, editors. Information systems: the challenges of theory and practice. Las Vegas, USA: Information Institute; 2003. p. 335–65.

Clark-Dickson P. Alarmed and dangerous; 2001 [e-Access March 2001].

Connolly P. Security starts from within. InfoWorld 2000;22(28): 39–40.

Detert J, Schroeder R, Mauriel J. A framework for linking culture and improvement initiatives in organisations. The Academy of Management Review 2000;25(4):850–63.

Eisenberger R, Cameron J. The detrimental effects of reward: myth or reality? American Psychologist 1996;51:1153–66.

Freeman E. E-merging risks: operational issues and solutions in a cyberage. Risk Management 2000;47(7):12–5.

Hartley B. Ensure the security of your corporate systems (developing a security policy). E-Business Advisor 1998;16(6):30–2.

Klein A, Masi R, Weidner C. Organisation culture, distribution, and amount of control, and perceptions of quality. Group and Organisation Management 1995;20:122–48.

Knapp KJ, Marshall TE, Rainer RK, Ford FN. Information security: management's effect on culture and policy. Information Management & Computer Security 2006;14(1):24–36.

Koh K, Ruighaver AB, Maynard S, Ahmad A. Security governance: its impact on security culture. In: Proceedings of the third Australian information security management conference, Perth, Australia; September 2005.

Lau O. The ten commandments of security. Computers and Security 1998;17(2):119–23.

Maynard S, Ruighaver, AB. What makes a good information security policy: a preliminary framework for evaluating security policy quality. In: Proceedings of the fifth annual security conference, Las Vegas, Nevada USA; 19–20 April 2006.

Mowday RT, Sutton RI. Organisational behavior: linking individuals and groups to organisational contexts. Annual Review of Psychology 1993;44:195–230.

Ngo L, Zhou W, Warren M. Understanding transition towards organisational culture change. In: Proceedings of the third Australian information security management conference, Perth, Australia; September 2005.

Nosworthy J. Implementing information security in the 21st Century – do you have the balancing factors? Computers and Security 2000;19(4):337–47.

Pierce WD, Cameron J, Banko KM, So S. Positive effects of rewards and performance standards on intrinsic motivation. The Psychological Record 2003;53:561–79.

Schein E. Organisational culture and leadership. 2nd ed. San Francisco: Jossey-Bass; 1992.

Schlienger T, Teufel S. Information security culture – the socio-cultural dimension in information security management. In: IFIP TC11 international conference on information security, Cairo, Egypt; 7–9 May 2002.

Schlienger T, Teufel S. Analyzing information security culture: increased trust by an appropriate information security culture. In: 14th International workshop on database and expert systems applications (DEXA'03), Prague, Czech Republic; 2003.

Schwarzwalder R. Intranet security. Database and Network Journal 1999;22(2):58–62.

Shedden P, Ahmad A, Ruighaver AB. Risk management standards– the perception of ease of use. In: Proceedings of the fifth annual security conference, Las Vegas, Nevada, USA; 19–20 April 2006.

Shinn MT. Security for your e-business. Enterprise Systems Journal 2000;15(8):18.

Tan TCC, Ruighaver AB, Ahmad A. Incident handling: where the need for planning is often not recognised. In: Proceedings of the first Australian computer network, information & forensics conference, Perth, 24 November 2003.

Von Solms B. Information security – the third wave? Computers and Security 2000;19(7):615–20.

Wood C. Integrated approach includes information security. Security 2000;37(2):43–4.

Tobias Ruighaver is a Senior Lecturer and Head of the Organisational Information Security Group in the Department of Information Systems at the University of Melbourne. His research interests are in the areas of Intrusion Detection, Forensic Investigations, Information Security Risk Assessment, Security Governance, and Security Culture.

Sean Maynard is a lecturer in the Department of Information Systems at the University of Melbourne. His primary research areas are in the area of information systems security, in particular focusing on the Evaluation of Security Policy Quality and on the investigation of Security Culture within organisations. He has also conducted research into the Evaluation of Decision Support Systems, and on early research in the use of computing technology to aid senior management (EIS).

Shanton Chang is a lecturer in Change Management and Social Impacts of Information Systems at the Department of Information Systems, University of Melbourne. He completed his Ph.D. in Managing Multicultural Workforces at Monash University. His current primary areas of research include the Social Aspects of Broadband Technology Adoption and Appropriation, Online Behaviour and Online Prosumers, and the Relationship between Cultures and Information Technology.

Computers & Security

A video game for cyber security training and awareness

Benjamin D. Cone, Cynthia E. Irvine, Michael F. Thompson, Thuy D. Nguyen*

Department of Computer Science, Center for Information Systems Security Studies and Research, Code CS/Ic, Naval Postgraduate School, Monterey, CA 93943, USA

Keywords:
Information assurance
Training and awareness
Educational simulation
Video game
Network security

ABSTRACT

Although many of the concepts included in cyber security awareness training are universal, such training often must be tailored to address the policies and requirements of a particular organization. In addition, many forms of training fail because they are rote and do not require users to think about and apply security concepts. A flexible, highly interactive video game, CyberCIEGE, is described as a security awareness tool that can support organizational security training objectives while engaging typical users in an engaging security adventure. The game is now being successfully utilized for information assurance education and training by a variety of organizations. Preliminary results indicate the game can also be an effective addition to basic information awareness training programs for general computer users (e.g., annual awareness training.)

1. Introduction

Typical employees of both large and small organizations may be made acutely aware of a wide array of cyber security problems. These range from spam and phishing to well organized attacks intended to corrupt or disable systems. Despite these constant reminders, users often take an ostrich-like attitude toward the security of the information systems they use, believing that there is little that they can do to mitigate this onslaught of problems. Even within the major organizations, users select trivial passwords or think that, so long as they keep their machines within viewing distance, arbitrary hookups to unknown networks and to the Internet pose no threat. Thus, despite their increased awareness of security problems, users and administrators of systems continue to take few effective precautions. Yet, to achieve an adequate security posture, organizations must combat this user apathy with effective training and awareness programs. The enormity of the problem associated with effective user training and awareness is evident in that it was considered one of five areas of highest priority for action in a national plan for cyberspace security (EOP, 2003).

Human factor studies illustrating the need for user training and awareness are well documented, e.g. (Whalen, 2001). The concept of using games to support health, education, management, and other sectors has resulted in a high level of interest and activity (Prenski, 2001). The tacit knowledge gained by applying concepts in a virtual environment can significantly enhance student understanding.

A number of games have been developed involving protection of assets in cyberspace. Some teach information assurance concepts, e.g. CyberProtect (DoD, 1999), whereas others provide pure entertainment with no basis in information assurance principles or reality (Nexus, 2003). None have presented an engaging virtual world that combines the human and technical factors associated with an IT environment. In addition, these games are limited in the scope of information assurance topics covered. Short of going back to the creator for

* *Corresponding author.*
E-mail addresses: benjamin.cone@hotmail.com (B.D. Cone), irvine@nps.edu (C.E. Irvine), mfthomps@nps.edu (M.F. Thompson), tdnguyen@nps.edu (T.D. Nguyen).

a new version, there is no way to add new material to the game.

Effective user security awareness training can greatly enhance the information assurance posture of an organization (NIST, 1993). Yet holding a trainee's attention sufficiently long to impart a message is a considerable challenge, particularly when the training is mandated and the target audience views the topic as potentially mundane. Video games have been proposed as an engaging training vehicle (Prenski, 2001). Here we describe a video game-like tool called CyberCIEGE and how it was employed to develop security awareness training targeted for the requirements of a specific organization, and how this extensible tool can offer training and education for a range of target audiences. For this analysis, training for uniformed and civilian personnel associated with the U.S. Navy has been conducted.

CyberCIEGE is unique in that it is a highly extensible game for teaching information assurance concepts. It may be applied to a wide range of audiences having different levels of technical sophistication. It has its own language for creating new educational scenarios and is accompanied by tools and tutorials that help instructors develop customized scenarios.

We start with a review of commonly used training and awareness techniques, and follow with an overview of CyberCIEGE and a more detailed description of how scenarios for the game are constructed. At this point, it will be possible to examine policies for information assurance training and awareness of our target organization and then describe a targeted requirement analysis. How two CyberCIEGE scenarios, one for general awareness and the other for IT personnel, were created to fulfill organizational information assurance training and awareness requirements will follow. This work concludes by pointing to several new directions for further development of the CyberCIEGE educational tool.

2. Background

To provide a context for subsequent discussion of CyberCIEGE as a tool for user training and awareness in information assurance, it is useful to review both current training and awareness methods as well as provide an overview of CyberCIEGE.

2.1. Common current training and awareness techniques

Training and awareness is generally accomplished using one or a combination of several techniques described below.

Formal Training Sessions can be instructor-led, brown-bag seminars, or video sessions. Formal training in sessions facilitated by local information security personnel represents the traditional approach to user training and awareness within the Department of the Navy. The success of this approach depends upon the ability of the training facilitator to engage the audience.

Passive computer-based and web-based training represents a centralized approach to the training and awareness problem. An example is the web-based training in information assurance offered by the U.S. Department of Defense (DoD, 2006). CBT offers the user the flexibility of self-paced training, and provides the organization with the ability to train users to an enterprise-wide standard. Its disadvantage is that training and awareness becomes a monotonous slide show that fails to challenge the user and provides no dialogue for further elaboration. Often, users attempt to complete CBT sessions with minimal time or thought. The CBT developer must attempt to provide engaging instruction within the constraints of a passive medium.

Strategic placement of awareness messages seeks to raise the level of consciousness through the delivery of messages in the workplace. Some of the more common delivery methods include organizational newsletters and memos, email messages, posters, screen savers, and security labels, e.g. posters highlighting various cyber security risks (CCS, 2006).

Interactive computer-based training, such as a video game, generally falls into two broad classes: first-person interaction games or resource management simulations. The majority of games falls into the first category and include first-person shooter games where the player is confronted by an adversary or problem and must take an appropriate action or is penalized, sometimes severely. In contrast, resource management games require the player to manage a virtual environment using limited resources. The player attempts to make choices that improve the environment within the constraints of the available resources. Good choices result in a richer environment and additional resources. SimCity™, other "sims" games, and RollerCoaster Tycoon (R) are popular examples of resource management games.

Games and simulations have become increasingly accepted as having enormous potential as powerful teaching tools that may result in an "instructional revolution" (Foreman, 2004). Prenski (2001) and Gee (2005) have provided a framework to construct and analyze games in education. The latter has described the context of a specific game as a semiotic domain that allows students to learn an area through use and experience while leading the student to approach problem solving through critical thinking. Analysis of the effectiveness of games is in its infancy; however, pioneering work (Gee, 2003; Aguilera and Mendiz, 2003; Squire, 2005; Gredler, 2004) is beginning to show that games offer an effective alternative to, or supplement for, more traditional modes of education. For example, through the use of virtual worlds, games provide a concrete experience within which students can internalize domain-specific concepts. Student's critical thinking skills are honed. In addition, the game format often appeals to students with short attention spans.

2.2. CyberCIEGE

In 2005, the Naval Postgraduate School released an U.S. Government version of CyberCIEGE, a video game intended to support education and training in computer and network security. Simultaneously, our collaborators at Rivermind, Inc. made a version available to non-government organizations. The game employs resource management and simulation to illustrate information assurance concepts for education and training (Irvine and Thompson, 2003, 2004). In the CyberCIEGE virtual world, players construct and configure the computer networks necessary to allow virtual users to be productive and achieve goals to further the success of the enterprise. Players operate and defend their networks, and can watch the consequences of their choices, while under attack by hackers, vandals and potentially well-motivated professionals.

2.2.1. *CyberCIEGE components*

The building blocks of CyberCIEGE consist of several elements: a unique simulation engine, a domain-specific scenario definition language, a scenario development tool, and a video-enhanced encyclopedia (Irvine et al., March 2005). CyberCIEGE is intended to be extensible in that new CyberCIEGE scenarios tailored to specific audiences and topics are easily created (Irvine et al., June 2005).

The scenario definition language expresses security-related risk management trade-offs for different scenarios. The CyberCIEGE simulation engine interprets this scenario definition language and presents the player with the resulting simulation. What the player experiences and the consequences of the player choices are a function of the scenario as expressed using the scenario definition language.

The game engine and the language that feeds it are rich in information assurance concepts so that it is possible to simulate sophisticated environments subject to a variety of threats and vulnerabilities. They also include substantial support for relatively brief, scripted training and awareness scenarios. This support includes cartoon-like balloon speech by the virtual users, message tickers, pop-up quizzes and conditional play of video sequences, e.g., a computer worm.

So that educators are able to assess their students, CyberCIEGE also produces a log of player activity. Triggers within the scenario cause output to be appended to the log where a number of status indicators may be recorded. A separate log is maintained for each player, thus allowing the instructor to track the progress of individual students.

The set of CyberCIEGE components is illustrated in Fig. 1.

2.2.2. *Development and Testing of CyberCIEGE*

The collaborative development of CyberCIEGE was an iterative process by video game developers with little knowledge of information assurance, and information assurance technologists with little background in video games. Early focus was on establishing a language that would allow us to construct scenarios as per our broad teaching objectives (Irvine and Thompson, 2003). This scenario development language evolved around a core ability to express a security policy in terms of users and information (Irvine and Thompson, 2004). The game developers built the CyberCIEGE game engine using C++ and their 3D graphics library. The engine was designed to consume the scenario definition language and behave in a logically consistent manner.

Some scenario language elements were relatively straightforward to represent in the game engine. For example, the costs of purchasing computer equipment, or the penalties incurred when users were not able to achieve goals are conceptually straightforward, and have analogues in other resource management games. Innovation was required to automate assessment of vulnerabilities in networks constructed by players such that the game could mount credible attacks. The attack logic within the game engine required several iterations and considerable testing. Dozens of test scenarios were generated to exercise the engine's response to a range of topologies, component configurations and attacker motives. Ultimately, some of the attack logic within the game engine was built around the tests rather than to any precise specification. For most scenarios, this has proven adequate. The resulting attack engine represents a range of security policies, including those modeled by Bell and LaPadula (1975) and Biba (1977), as well as discretionary security policies (Lunt, 1989; Bishop, 2002).

CyberCEIGE has been the basis for several master theses at NPS in which students developed their own scenarios. This

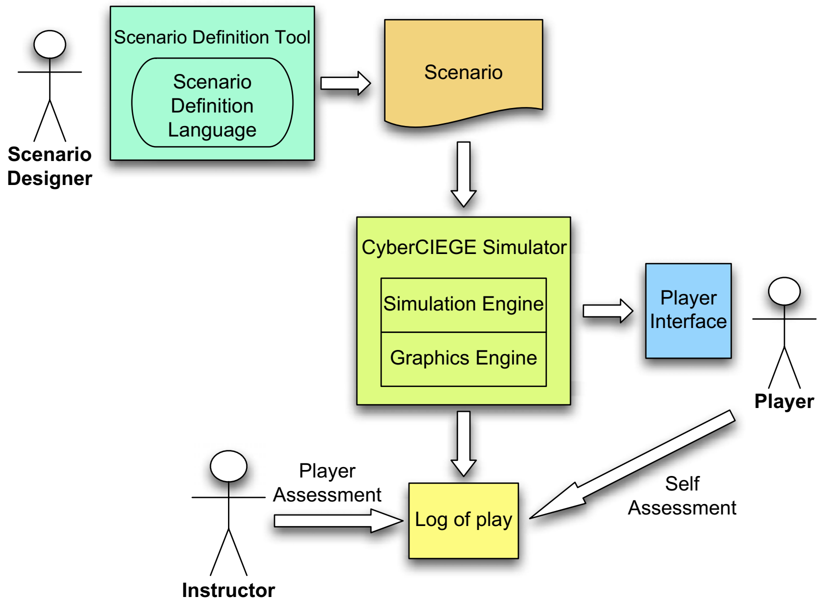


Fig. 1 – CyberCIEGE components.

early exercising of the CyberCIEGE engine provided considerable breadth and depth of informal testing. The CyberCIEGE user interface has undergone more formal testing by the Pacific Northwest National Laboratories (Roberts et al., 2006). That testing resulted in several refinements of the game interface.

The scenario development tool (Johns, 2004) and the related tools (Teo, 2003) were developed using Java. And again, student theses work provided substantial informal testing of these tools.

The game and the tools all are designed to run on the Windows 2000 and Windows XP operating system. The graphic libraries make considerable use of the DirectX interface to render the three dimensional world.

3. Scenario construction

Story telling is the key to a good CyberCIEGE scenario. The player should readily grasp the nature of the virtual environment (e.g., a small business with valuable intellectual property) and the types of choices that he has to make. Within this context, the player should have reason to care about the ramifications of these choices.

Scenario designers utilize the scenario definition language to construct a virtual environment that drives players to make resource management decisions. These player choices affect the productivity of an enterprise, and affect the vulnerability of information assets to compromise by virtual attackers. The CyberCIEGE game engine interprets the scenario definition language, presenting the player with a virtual environment as defined by the designer. To construct a scenario, the designer must understand the semantics of the scenario definition language and the capabilities of the CyberCIEGE game engine. A form-based integrated development environment allows designers to construct scenarios without mastering the syntax of the design language (Johns, 2004).

3.1. *Programming the CyberCIEGE game engine*

In every CyberCIEGE scenario, the player is the information assurance decision maker for some enterprise. An enterprise may be a large military facility, or it may be a home office. The fundamental abstractions within the CyberCIEGE game engine are not computers, networks and protection mechanisms. Rather, they are assets, users, and attackers (Irvine and Thompson, 2003). Assets are information resources. Users are typically employees of the enterprise who have goals that require computerized access to assets. Players succeed by facilitating user access to assets. Some assets have substantial value to the enterprise based on secrecy or integrity. And some assets may have value based on their availability. Assets also have value to attackers, and this motive determines the means by which the attacker will attempt to compromise an asset. Player choices affect the opportunity (or lack thereof) for the attacker to compromise the assets. The enterprise (and by extension the player) is penalized the value of an asset should it be compromised or made unavailable.

Within any given scenario, the users, assets, and attackers are for the most part fixed by the designer and are not modified by player choices. Designers also specify the initial state of the scenario (e.g., an initial set of computers) and dynamic changes to the scenario (e.g., the introduction of new user goals.)

Players see the enterprise as an animated three dimensional representation of an office building or military headquarters. Each scenario has one main office and an optional small offsite office. Users inhabit these buildings, wandering about or productively sitting at desks in front of computers. If computers are available, either as a scenario default or through purchase by the player, users will create and access assets using the computers. This user behavior is driven by the user goals specified by the designer. If computers are networked together, users may access assets over the network. Network devices such as routers enable users to access the Internet, and allow attackers on the Internet to potentially access enterprise resources. Suitably motivated attackers can enter buildings to compromise assets. They may compromise computer-based protection mechanisms, and may wiretap network links. Attackers may also bribe untrustworthy users to compromise assets. Finally, users themselves may have motive to compromise assets.

Players may hire guards to help with the physical protection of buildings or offices within buildings. Players may purchase physical protection mechanisms such as alarms and they may select which users are permitted to access different physical areas (i.e., "zones") within the virtual buildings. Procedural security choices affect user behavior (e.g., leaving computers logged in). Players can purchase user training to improve user adherence to procedural policies.

3.2. *The game engine as illustrated with a simple scenario*

Consider a scenario consisting of a single asset and a single user having a goal to read the asset. If this scenario is fed to the CyberCIEGE engine, the user will fail to achieve the goal of reading the asset until the player buys the user a computer. The designer associates a productivity value with the user that affects the size of the penalty resulting from failure to access the asset. When the player purchases a computer, the user will create the asset on the computer. Once it exists on a computer, attackers potentially target the asset. For example, an attacker might break into the office housing the computer and walk off with the entire computer. Or, if the asset's attacker motive is based on integrity, the attacker might hack into the computer and modify the data. If the asset is compromised, the player is penalized as specified when the designer defines the asset.

In the above example, the designer simply defines a user and an asset. The game engine manages the rest (Irvine and Thompson, 2004). The game engine manages the virtual economy to reward players when users are productive and to penalize them when goals are not achieved or assets are compromised. It also includes a sophisticated attack engine to assess the suitability of the current protection mechanisms to protect the assets based on the asset motive.

3.3. *Extending the simple scenario*

In addition to defining assets and users, the designer specifies the initial state of the scenario including:

- Physical security properties (e.g., alarms, guards, etc.) of the different zones (e.g., offices);
- The set of pre-existing computers and their configurations including network connections;
- Procedural security policies to be followed by the users;
- Initial user training (This affects adherence to procedural policies.);
- Background checks for different kinds of users (e.g., based on user clearance);
- Which kinds of attacks will be initially active and which will be suppressed;
- How much money the player will start with;
- The kinds of computers and network devices available for purchase; and
- Support staff available to help, administer and maintain computer systems.

3.4. *Interacting with the player and dynamically altering the scenario*

The CyberCIEGE scenario definition language allows scenario designers to periodically assess the ongoing game state ''conditions'' and respond using active ''triggers''. Game state conditions include such things as the passing of time, whether users are achieving their goals, computer configuration settings and whether attackers have compromised assets. Active triggers include pop-up messages, brief movies, changes in user goals, commencement of attacks, and user feedback to the player via balloon speech as reflected in Fig. 2.

Scenarios are divided into multiple phases, each of which includes one or more objectives that the player must achieve prior to moving on to the next phase. Designers use conditions and triggers to assess whether objectives have been met and to change the environment for the next phase (e.g., introduce additional user goals).

3.5. *Scenario audience selection*

The first step in the design of a scenario is to identify its purpose and audience. For example, does the intended audience have experience with computer games? Are they expected to have played other CyberCIEGE scenarios and thus have some level of mastery of the mechanics of the game?

The scenario definition language supports a broad range of different types of scenarios. At one end of the spectrum are simple scripted scenarios such as those intended for basic training and awareness. These scenarios are designed to affect user behavior where human factors are the sources for potential security compromises (Whalen, 2001), e.g., ''beware of email attachments.'' This type of scenario is often built entirely from conditions and triggers, with very little reliance on the game engines' economy or attack engines. For example, a set of conditions assess whether the user has been instructed to beware of email attachments, and triggers provide direct feedback based on that game state. At the other end are sophisticated scenarios for players who have a basic understanding of network security engineering. These scenarios rely more on the game engine itself to direct attacks and manage the overall economy.

3.6. *Elements of scenario design*

The scenario designer defines the information assets. What kind of information is it? What is the asset value and what makes it valuable? Why would an attacker target the asset? The designer also defines the users. What assets do the users need to access? Why do they need to access them? Do users need to share assets? Do users require access to assets via the Internet (e.g., publicly available documents)?

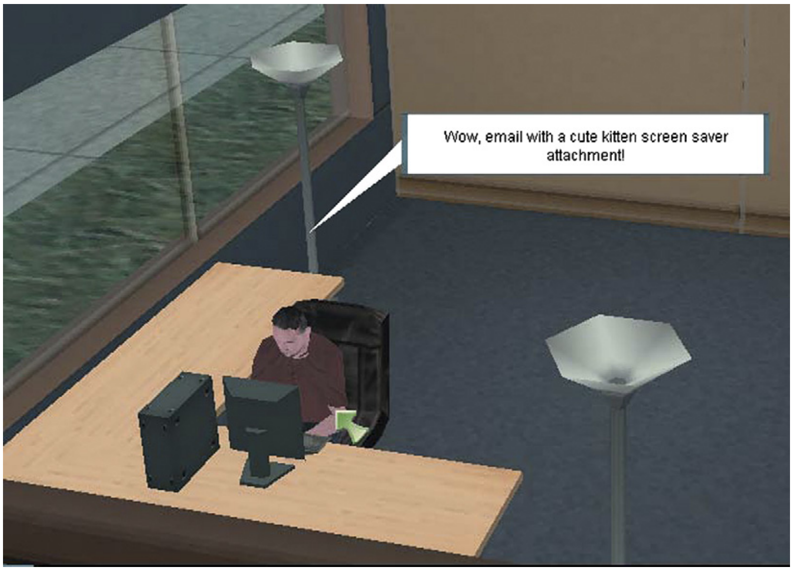


Fig. 2 – Pop-up messages can be initiated using active triggers.

The scenario designer describes the story line in the scenario briefing and in the descriptions of the assets and users. This textual information is intended to provide players with the context of the scenario. The designer describes individual player objectives and specifies the conditions that constitute the achievement of each objective. The initial state of the scenario can be used to constrain a player's options. For example, a player can be given insufficient cash to fully secure a site until an initial set of objectives is achieved.

Finally, the designer specifies feedback to move the player along through the scenario based on current game conditions. For example, what should a user say or think if a specific goal cannot be met? The engine causes the user to wander aimlessly or violently pound on the keyboard. The designer can enhance this with specific user ''thoughts'' or comments that appear in bubble text. In some scenarios the designer may choose to assess the suitability of protection mechanisms using conditions and warn the player prior to the attack engine's exploitation of the vulnerability. And in other scenarios the designer will provide substantial help tips to aid the player with the mechanics of the tool. CyberCIEGE includes a rich on-line encyclopedia that can serve as context-dependent help. Ultimately the designer selects the conditions that constitute a ''win'' or a ''loss'', and provides the text to display in the respective debriefing. The encyclopedia includes several animated movie tutorials (e.g., describing malicious software) that can be launched as a part of the debriefing.

3.7. *Integrated development environment*

Designers build and modify scenarios using the *Scenario Development Tool* (SDT), which automates the syntax of the CyberCIEGE scenario definition language through the use of reusable libraries and forms having pull down menus (Johns, 2004). As is illustrated in Fig. 3, the SDT permits the designer to compile and run scenarios, and then view a formatted presentation of the resulting log (Teo, 2003). In Figs. 3 and 4 the numbered arrows indicate the sequence of interactions. The SDT performs input validation and limited consistency checking (e.g., ensuring that references to users are valid). The SDT Users Guide (CISR, 2002) includes a tutorial that walks new designers through the construction of a complete scenario. The SDT was used to construct the scenarios described below. The source for these and other scenarios is distributed with CyberCIEGE, and developers may use the SDT to alter or expand the scenarios. Upon completing development or revision of a scenario, designers use the Campaign Manager to group the scenario with other scenarios into a collection of scenarios that are to be played by students in sequence as illustrated in Fig. 4. Instructors can view summaries and details of student progress via the Campaign Analyzer as illustrated in Fig. 5.

At this point, it is possible to examine how CyberCIEGE can be applied to the training and awareness needs of a real organization, in our example, the U.S. Navy.

4. Requirements elicitation

Two factors determine the requirements for the use of CyberCIEGE as a training and awareness tool in the context of the U.S. Navy. The first is the collection of policies that mandate training and awareness activities within the Military and the Navy. This is followed by an analysis of specific topics that must be addressed.

4.1. *Current policies for IA training and awareness*

Like the other services, the U.S. Navy must adhere to laws and directives intended to cover the entire Department of Defense (DoD). This section will describe the important laws and policies that have affected Navy choices with respect to training and awareness in information assurance.

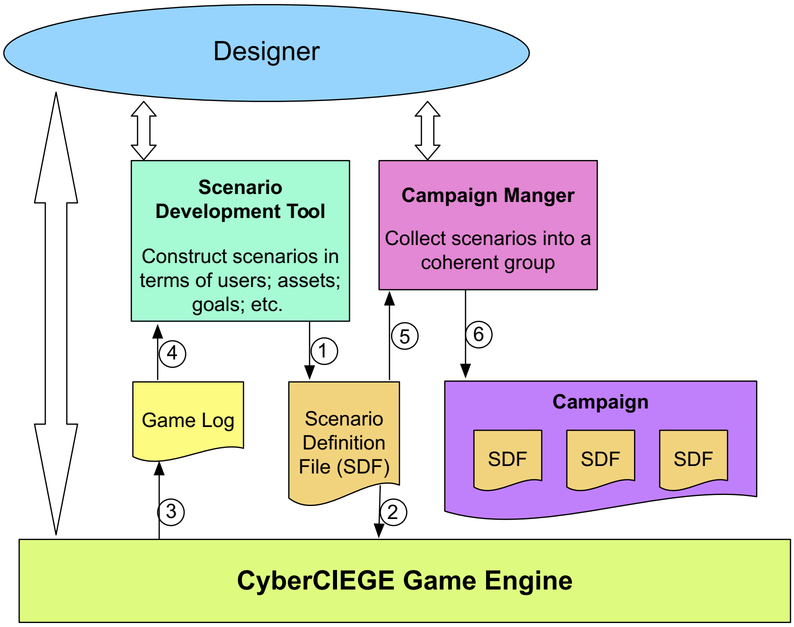

Designer
Scenario Development Tool
Construct scenarios in terms of users; assets; goals; etc.
Campaign Manger
Collect scenarios into a coherent group
1
2
3
4
5
6
Game Log
Scenario Definition File (SDF)
Campaign
SDF
SDF
SDF
CyberCIEGE Game Engine

Fig. 3 – Scenario designers use the SDT and the Campaign Manager.

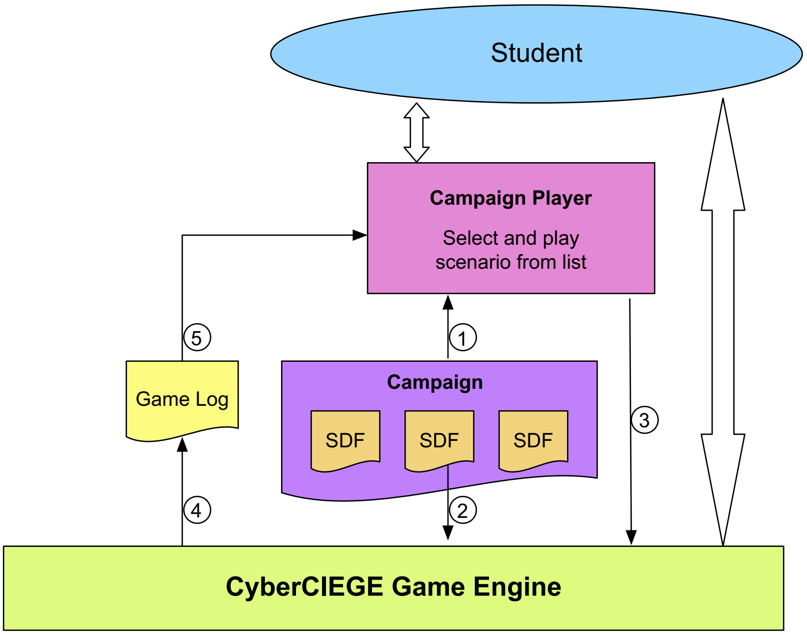


Fig. 4 – Students use the Campaign Player.

The United States Computer Security Act of 1987 mandated periodic security training for all users of Federal information systems. In response, the Department of the Navy placed the burden of responsibility for training and awareness on local Information Systems Security Managers (NSOP, 1995), who were, in turn, responsible for developing local training sessions or CBT. To supplement other IA directives (DoD, 2002, 2006), in 2004, the U.S. Department of Defense issued DoD Directive 8570.1 (DoD, 2004), which mandated initial and annual refresher information assurance training for all DoD information system users. Since then, all users of Navy information systems have been instructed to complete a DoD IA awareness CBT. The CBT is a web-enabled slide presentation. It is trivial for a personnel to click through the training to its successful completion without absorbing any of the material.

Directive 8750.1 has highlighted the importance of fostering a security culture and the need to find training techniques that will actively engage the typical user. A participatory video game requires more user involvement than slide presentations or other standard training and awareness vehicles.

4.2. *Requirements analysis*

Training and awareness requirements were developed from the legacy Information Security program of the U.S. Navy and from the current Department of Defense IA training and awareness computer-based training course.

Many of the requirements for the awareness scenario were obtained from the U.S. Navy Information Security Program. Navy requirements for user security training are found in the Navy INFOSEC program guidebooks for local Information System Security Officers (NSOP, February 1996) and Network Security Officers (NSOP, March 1996). These documents offer recommended training curriculum topics and subtopics including:

- The value of information, e.g., personnel files, legal records, and trade secrets
- Communication and computer vulnerabilities such as malicious software, Internet risks, human errors, and Internet security risks
- Basic safe computing practices such as locking computers when unattended
- Password management including password generation, protection, and change frequency
- Local security procedures, e.g., cipher locks and violation reports

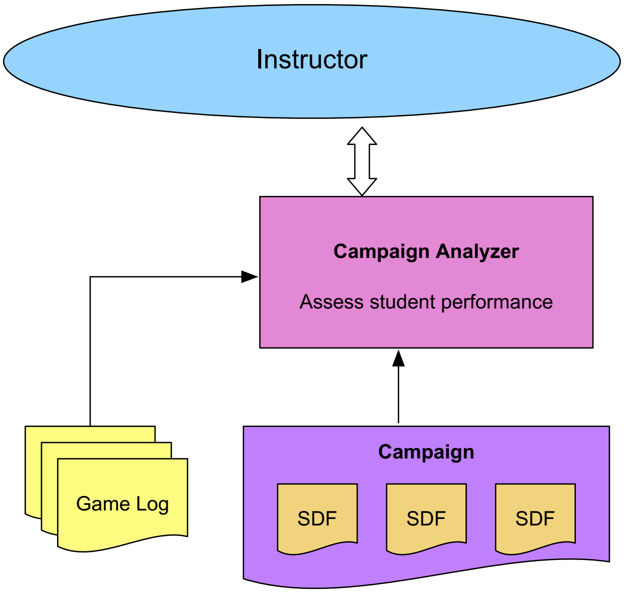


Fig. 5 – Instructors use the Campaign Analyzer.

The other requirements source was the DoD Information Assurance Awareness CBT. The majority of naval organizations currently use the ‘‘DoD Information Assurance Awareness’’ CBT (DoD, 2006) to fulfill obligations for enterprise-wide annual refresher training. It addresses the following topic areas:

- Importance of IA (overview, evolution, and policy)
- IA threats (threats, vulnerabilities, social engineering, and internet security)
- Malicious code (overview, protection, and internet hoaxes)
- User roles (system security and protecting DoD information)
- Personal and home security (on-line transactions and security tips)

These topics provided the requirements for the video game-based training and awareness.

5. Scenarios for training and awareness

Two CyberCIEGE scenarios were designed to fulfill the Navy IA training requirements. The first seeks to make the player aware of basic IA problems and principles. The second is intended for more sophisticated users of computer-based assets. A brief summary of other CyberCIEGE awareness and training scenarios is provided in Section 5.2.

The basic user scenario focuses on computer security fundamentals. The player is placed in the role of a security decision maker aboard a ship, who must complete objectives that raise the security posture of the organization. If objectives are not completed within a specified time, appropriate attacks are triggered by the game engine and the player is penalized. After completing each objective, the player is presented with an awareness message that relates the action taken in the game with real-life circumstances and provides feedback regarding the players choices. The player wins by completing all the objectives without incurring ‘‘fatal’’ penalties.

For each topic identified in the requirements analysis, a scenario element was created that requires the player to do something that will convey the concept to be learned. Some of the topics and activities are described in Table 1. Features that made this scenario Navy-specific included the protection of classified information and cultural aspects of organizational security associated with the hierarchical command structure of the DoD.

Table 1 – Basic awareness topics and player activities

Topic	Player activity
Introductory IA briefing	This briefing includes definitions and descriptions of important IA elements and how they interact.
Information value	The user must protect high value information and answer questions about information dissemination.
Access control mechanisms	The player is introduced to both mandatory and discretionary access control, with the latter as a supplement to controls on classified information.
Social engineering	The player is presented with a scenario that will lead to a social engineering attack if proper action is not taken.
Password management	The player must prevent a game character from revealing his password to an outside contractor.
Malicious software and basic safe computing	The player must determine and expend resources to procure three procedural settings that will prevent malicious software propagation.
Safeguarding data	The player is presented with a situation where it appears that a game character is leaving the premises with sensitive information. Actions taken by the player allow the importance of secure storage of backups to be conveyed.
Physical security mechanisms	The player must select cost-effective physical security mechanisms to prevent unauthorized entry into sensitive areas.

5.1. *Scenarios for IT staff*

Navy IT training requirements for staff with IT-related jobs are addressed by a second scenario that focuses on network security, and serves to introduce technical users into the roles they must assume. The player assumes the role of acting security manager while the ‘‘boss’’ is away. The player must manage three internal networks, one of which processes classified information. During this scenario, the player must complete technical objectives addressing physical security mechanisms, access control, filtering, antivirus protection, data backups, patching configurations, password policies, and network vulnerability assessment.

5.2. *Other scenarios*

The rich and flexible CyberCIEGE scenario definition language supports information assurance training beyond military environments. For example, an ‘‘identity theft’’ scenario was built to teach users about the methods of identity theft prevention in home computing environments (Ruppar, 2005). This scenario focuses on a few basic user behaviors that can greatly reduce the risk of identity theft, while highlighting consequences of risky behavior through an engaging story line.

One set of scenarios was developed solely to help train users to reduce the risks of distributing worms and viruses. Here, the player can see the damaging effects of worms and viruses, and learns that a major cause of malicious software proliferation is through user execution of email attachments.

Other CyberCIEGE scenarios illustrate more complex and subtle information assurance concepts. These longer, more sophisticated scenarios are more like traditional simulation and resource management games. For these, the target audience may be advanced computer security students, or information security decision makers.

Several students have developed relatively complex scenarios as part of their master’s thesis work, an example of which is described by Fielk (2004). And while not all such efforts have resulted in polished games that are fun to play, the process of building scenarios requires students to confront fundamental information assurance issues in order to build a consistent virtual environment. For example building a scenario requires the student to explain an asset's value to the player in a way that the player can understand both the consequences of asset compromise, and the motives of would-be attackers.

The two Navy IA training scenarios described above were completed as part of a master's thesis. Development of a basic scenario, including a substantial learning curve for the SDT, requires between 330 and 400 h of work depending on the student's aptitude and programming skills.

6. Discussion and future work

This paper demonstrates that information assurance awareness and training can be provided in an engaging format. CyberCIEGE was employed to meet a specific set of Navy IA training requirements, thus demonstrating that it is sufficiently flexible to illustrate a range of security topics in a variety of environments, both generic and organization-specific. Initial test results for the basic user training scenario are positive and illustrate the utility of CyberCIEGE in supporting awareness programs.

6.1. *User experiences*

CyberCIEGE was originally developed to be made available at no cost to organizations of the federal government of the United States. Since then, our development partner elected to also make it available at no cost to schools and universities. To date, approximately 130 organizations have made inquires at the CyberCIEGE website (CISR, 2006) and have been given download instructions. A number of these organizations currently use the game as a training tool.

The tool is used at our institution within our information assurance curriculum, and has been the subject of several master theses as described in Section 2.2.2.

These and more casual user experiences have resulted in feedback on CyberCIEGE, which has led to a number of recent improvements.

6.2. *Future work*

The effectiveness of CyberCIEGE for basic information assurance awareness has not yet been fully assessed. While initial feedback has been positive, a side-by-side comparison with traditional on-line click-through awareness programs (DoD, 2006) is needed. This testing would include a test group that only recieves CyberCIEGE training, one group that only receives click-through training and one group that receives both. Our informal experiences show that some users simply will not expend any effort to learn even the most basic mechanics of a video game. For these users, interactive training methods will not be effective if they require anything more involved that the repeated clicking of a mouse or pressing of an enter key. On the other hand, those users with some experience in video games or adventure games appear more inclined to explore the game, sometimes proceeding beyond the simple awareness scenarios into more sophisticated scenarios. A test study with a relatively large user pool would help quantify the usefulness of CyberCEIGE in place of or in addition to existing on-line awareness programs.

There are several functional aspects of CyberCIEGE for which future work is planned. First, it would be useful for instructors to be able to monitor the ongoing progress of students as they advance through either a single scenario or a campaign of several scenarios. Additional mechanisms and tools will be required in the CyberCIEGE framework to support this capability.

The ability of the scenario designer to use triggers and other dynamic mechanisms to cause changes in the evolution of a scenario is one of the greatest strengths of CyberCIEGE. Further investigation is required to determine additional techniques to introduce dynamic content in the game. In addition, the tool would benefit from better development interfaces with which to experiment with and test dynamic content.

Many video games involve multiple users and such activity is envisioned for CyberCIEGE. We have conducted a requirements analysis for a multiplayer version of CyberCIEGE and have determined how best to engage multiple players without turning it into an exercise that would give the appearance of promoting misbehavior on IT systems. Players are assumed to be concerned about partners with whom they might conduct cyber-based business interactions. To determine whether other systems are qualified to be a participant in the protection of his information assets, a player would conduct various tests on these foreign systems. The game would consist of a scenario-specific number of rounds of preparation and testing by all nodes. As with existing single-player scenarios, tests could be focused on a particular information assurance issue, such as passwords or firewall configuration, or could cover a broad range of topics.

CyberCIEGE is currently designed to address wired networks. A more advanced version of the game could include both wired and wireless platforms. For the latter, issues associated with user and platform mobility, platform resources, wireless authentication, etc. could be addressed. In addition, CyberCIEGE could anticipate the security challenges that will be encountered in the next generation of processors. These include the management of virtual machine monitors and their guest operating systems in virtual machines, platform monitoring and attestation, distributed system management, and the balance between corporate convenience and individual privacy.

Acknowledgments

This work has been supported by the United States Navy, by the Department of Defense, and by the Office of Naval Research. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views the US Navy, Department of Defense, or the Office of Naval Research.

REFERENCES

de Aguilera M, Mendiz A. Video games and education: (education in the face of a ''parallel school''). Computers in Entertainment 2003;1(1):1–10.

Bell DE, LaPadula L. Secure computer system unified exposition and multics interpretation. Technical Report ESD-TR-75-306. Hanscom AFB, MA: MITRE Corp.; 1975.

Biba KJ. Integrity considerations for secure computer systems, Technical Report ESD-TR-76-372. MITRE Corp.; 1977.

Bishop M. Computer security: art and science. Reading, Massachusetts: Addison-Wesley; 2002.

CISR. CyberCIEGE scenario development tool users guide march. Monterey, CA: Naval Postgraduate School; March 2006.

CISR. CyberCIEGE, <http://cisr.nps.edu/cyberciege/>; 2002.

Central Coast Security. Security awareness/motivation posters, <http://members.impulse.net/~sate/posters.html#CompuSec>; September 2006 [Last accessed 11 September 2006].

DoD Directive 8500.1. Information assurance; 24 October 2002. Available from: <http://www.dtic.mil/whs/directives/corres/html/85001.htm> [Last accessed 20 June 2006].

DoD Directive 8570.1. Information assurance training, certification, and workforce management; 15 August 2004. Available from: <http://www.dtic.mil/whs/directives/corres/html/85701.htm> [Last accessed 20 June 2006].

Executive Office of the President. The national strategy to secure cyberspace. Available from: <http://www.whitehouse.gov/pcipb/>; 2003 [Last accessed 15 September 2006].

Fielk KW, CyberCIEGE scenario illustrating integrity risks to a military-like facility. Masters thesis. Monterey, CA: Naval Postgraduate School; September 2004.

Foreman J. Video game studies and the emerging instructional revolution. Innovate(1), http://www.innovateonline.info/, 2004;1 [Last accessed May 2006].

Gee JP. What video games have to teach us about learning and literacy. Plagrave Macmillan; 2003.

Gee JP. What would a state of the art instructional video game look like? Innovate(6), http://www.innovateonline.info/, 2005; 1 [Last accessed May 2006].

Gredler ME. Games and simulations and their relationships to learning. In: Handbook of research on educational communications and technology. 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 2004. p. 571–81.

Irvine CE, Thompson MF. Teaching objectives of a simulation game for computer security. In: Proceedings of informing science and information technology joint conference, Pori, Finland; June 2003. p. 779–791.

Irvine CE, Thompson MF. Expressing an information security policy within a security simulation game. In: Proceedings of the sixth workshop on education in computer security. Monterey, CA: Naval Postgraduate School; July 2004. p. 43–9.

Irvine CE, Thompson MF, Allen K. CyberCIEGE: an information assurance teaching tool for training and awareness. In: Federal information systems security educators' association conference, North Bethesda, MD;March 2005.

Irvine CE, Thompson MF, Allen K. CyberCIEGE: an extensible tool for information assurance education. In: Proceedings of ninth colloquium for information systems security education, Atlanta, GA; June 2005. p. 130–138.

Johns KW. Toward managing and automating CyberCIEGE scenario definition file creation. Masters thesis, Monterey, CA: Naval Postgraduate School; June 2004.

Lunt TF. Access control policies: some unanswered questions. Computers and Security 1989;8(1):43–54.

Nexus Interactive. AI Wars: the awakening, <http://www.aiwars.com>; 2003 [Last accessed November 2003].

National Institute of Standards and Technology. People: an important asset in computer security. NIST-CSL Bulletin October 1993.

Information systems security manager (ISSM) guidebook. Navy Staff Office Pub. 5239-04; 1995.

Information systems security officer (ISSO) guidebook. Navy Staff Office Pub. 5239-0; February 1996.

Network security officer (NSO) guidebook. Navy Staff Office Pub. 5239-08; March 1996.

Prenski M. Digital game-based learning. New York: McGraw-Hill; 2001.

Roberts IE, McColgin DW, Greitzer FL, Huston K. Usability and training effectiveness evaluation of CyberCIEGE. Pacific Northwest National Laboratory; January 2006.

Ruppar C. Identity theft prevention in CyberCIEGE. Masters thesis, Monterey, CA: Naval Postgraduate School; December 2005.

Squire K. Changing the game: what happens when video games enter the classroom? Innovate(6), http://www.innovateonline.info/, 2005;1 [Last accessed May 2006].

Teo TL. Scenario selection and student selection modules for CyberCIEGE. Masters thesis, Monterey, CA: Naval Postgraduate School; December 2003.

U.S. Department of Defense: Defense Information Systems Agency. CyberProtect, version 1.1. DoD IA training and awareness products, <http://iase.disa.mil/eta/prod-des.html> July 1999 [Last accessed 17 June 2006].

U.S. Department of Defense: Defense Information Systems Agency. DoD information assurance awareness course, <http://iase.disa.mil/dodiaa/launchPage.htm> February 2006 [Last accessed 11 September 2006].

Whalen T. Human factors in coast guard computer security – an analysis of the USCG's current awareness level and potential techniques to improve security program viability. Masters thesis, Monterey, CA: Naval Postgraduate School; June 2001.

Benjamin D. Cone is a Lieutenant in the United States Navy. He holds a B.S. from the United States Naval Academy, Annapolis, Maryland, and a M.S. in Information Technology Management from the Naval Postgraduate School in Monterey, California. Since graduation he has been at sea.

Cynthia E. Irvine is a Professor in the Department of Computer Science at the Naval Postgraduate School in Monterey, California. She holds a B.A. in physics from Rice University and a Ph.D. in astronomy from Case Western Reserve University, Cleveland, Ohio. Her fields of interest include inherently trustworthy systems, security architectures, and security education.

Michael F. Thompson is a Research Associate in the Center for Information Systems Security Studies and Research at the Naval Postgraduate School in Monterey, California. He also serves as Lead Security Engineer for Aesec Corporation. He holds a B.S. in Electrical Engineering from Marquette University. His research interests include security engineering and highly secure systems.

Thuy D. Nguyen is a Research Associate in the Department of Computer Science at the Naval Postgraduate School in Monterey, California. She holds a B.S. in Computer Science from the University of California at San Diego. Her research interests are high assurance systems, security engineering, and security requirements elicitation.

ELSEVIER

Computers & Security

Phishing for user security awareness

*Ronald C. Dodge Jr.**, *Curtis Carver*, *Aaron J. Ferguson*

Department of Electrical Engineering and Computer Science, United States Military Academy, West Point, NY, USA

Keywords:
Phishing
Email security
Social engineering
Security training
Information assurance
Spam
Computer user education

ABSTRACT

User security education and training is one of the most important aspects of an organizations security posture. Using security exercises to reinforce this aspect is frequently done by education and industry alike; however these exercises usually enlist willing participants. We have taken the concept of using an exercise and modified it in application to evaluate a users propensity to respond to email phishing attacks in an unannounced test. This paper describes the considerations in establishing and the process used to create and implement an evaluation of one aspect of our user information assurance education program. The evaluation takes the form of a exercise, where we send out a phishing styled email record the responses.

Published by Elsevier Ltd.

1. Introduction

The quest for information systems security has a significant, almost self cancelling facet—the user. User information assurance (IA) awareness is a random variable that is very difficult to characterize due to user's individual nature. Users create an open back door into our corporate networks through their internet enabled services, third party application use, and electronic interaction (i.e. email) with other users. This vulnerability is increased from mobile systems that join home and other commercial networks. While the application of host and network based security applications can provide some mitigation against malicious activity, there is no static technical defensive measure that can mitigate the threat introduced by user behavior. One of the most common interactions users have with entities outside control of our local networks is email. The July 2006 report issued by the Anti-Phishing Working Group reported 23,670 unique phishing attempts targeting over 14,191 websites used to commit identity theft, fraud and other malicious activity. These websites are very dynamic in nature, existing only for an average 4.8 days (Anti-Phishing Working Group, 2006). Security training and awareness programs have done a good job of mitigating this risk – but just how good? What measures exist to verify that users understand and consistently apply the best practices they are exposed to during periodic training?

The use of exercises to reinforce concepts in an educational setting has been written about frequently (Dodge et al., 2005). The United States Military Academy (USMA) has been very active in implementing hands-on exercises such as the Cyber Defense Exercise (Dodge et al., 2003). Typically, these exercises involve participation by knowing participants and involve a network attack/defense scenario. The United States Military Academy took the concept of an active learning and developed an email phishing exercise with the intent of evaluating the efficacy of our user IA training. The exercise first ran as a prototype in the spring of 2004 and has since been run two additional times. The most recent exercise (at the time of this writing) ended in November 2005.

The exercise was named Carronade after the Navy cannon used in the early 1770s. The inventors, Charles Gascoigne, Lt. General Robert Melville, and Patrick Miller, designed the

* *Corresponding author.*
E-mail addresses: ronald.dodge@usma.edu (R.C. Dodge), curtis.carver@usma.edu (C. Carver), aaron.ferguson@usma.edu (A.J. Ferguson).

cannon in 1759 while working at the Carron Iron Works Company on the Carron River in Stirlingshire, Scotland. The short cannon weighed about 68 pounds. They initially called it the "Smasher," but it was not adopted until 1779, and was then known as the Carronade. The Carronade although possessing limited range, was destructive at close quarters (less than 0.6 miles). It is important to note that in offensive operations during the 1700s, the objective was not to sink an enemy vessel but rather to avoid damaging the hull so as to capture it as intact as possible, so it would be retained as a "prize".

In keeping with this military theme, this exercise was named the Carronade because: (a) while the email had the potential to be destructive, the intent was to get the attention of cadets, not to cause damage to the Academy network or to penalize the cadets; and (b) the exercise was short range – conducted inside the USMA security perimeter – only cadets with a usma.edu domain name could launch the embedded link.

In this paper, we will present a background discussion on the exercise, describing its origin and planning considerations. We will further describe the evolution of the exercise from a prototype to a multi-email exercise designed to evaluate different forms of phishing and the efficacy of training. We will provide results from each exercise and offer some assessment of our awareness and training program. We then conclude with a look toward future exercises.

2. Background and previous work

West Point has two primary mechanisms for reaching out to our students with information assurance training and education. The first is through our curriculum for those students enrolled in our information assurance focused courses. The second is through a series of required training provided to every student. The curriculum leads to our capstone information course that is lab focused, providing students with practical exercises designed to reinforce information assurance principals. The course uses the Military Academy Cyber Defense Exercise as the capstone student project (Schepens and James, 2003). The general student population receives information assurance awareness training each semester through a student run program. The student body at USMA (called the corps of cadets) is organized in a military structure; broken down into four regiments, each with eight companies. Each company has approximately 130 cadets. Each company also has an Information Security Officer (ISO) who is responsible for conducting the IA training. In addition to the training sessions, the ISO also leads inspections of the student computers attached to the USMA network, ensuing patches are up to date, anti-spyware and anti-virus applications are functioning, and there are no unauthorized applications. This is known as the Information Technology Saturday Morning Inspection (IT-SAMI). This unique student organization facilitates the IA awareness program, however, is not required for the phishing exercise described in this paper.

As briefly discussed earlier, information assurance exercises are not new and specifically have been implemented at USMA for over 6 years. USMA has worked to enhance and improve the annual Cyber Defense Exercise (Schepens and James, 2003) and develop methodologies where the exercise could be implemented at other institutions (Dodge et al., 2005). Additionally, USMA has participated in the University of California, Santa Barbara sponsored Capture the Flag (CTF) exercise (Vigna, 2003). USMA is also implementing a local CTF that will be available only to teams on a local air gapped network and available in a virtual machine based format.

We feel our students who have participated in these cyber defense style exercises have a very high understanding of the impacts of poor user computer habits. However, the exercises only involve approximately 30 students (who happen to be predisposed to be technically aware). This leaves roughly 4000 students whose only exposure to information systems security is the USMA security awareness and training program. There exists no formal mechanism to evaluate the success of these programs. The development of the email phishing exercise was in direct response to a question of how well our user awareness programs work.

Shortly after briefing the email phishing exercises at several events, the New York state cyber security office implemented a similar exercise after consultation with USMA (Bank, 2005). The NY state exercise, and several others that were not named in the article, demonstrate the cross-agency utility of this exercise. The NY state exercise was developed in cooperation with the research director of the SANS institute, Alan Paller.

We begin explaining the exercise by first addressing the general considerations and then examining the specific goals common to each phishing exercise. We first must recognize that USMA is a very "connected" campus. In 2006, USMA was ranked as the 13th most "wireless" school in the United States (U.S. News). Each student has a laptop computer with a specific suite of software. While they are allowed to install third party applications that are not part of the official suite, they must purchase them on their own and the computer staff at USMA will not support them. Given these limitations virtually all students have the same basic configuration. In addition, email is a very heavily relied upon management and information dissemination tool.

2.1. General considerations

Depending on your organization, the specific implementation of an email phishing exercise might vary. However, certain general considerations are transparent and are important to address. The most important consideration is to define the objectives of the exercise. The exercise needs to have a clearly defined goal that is consistent with the awareness and training program in place.

A second consideration is to devise a mechanism for assessing IA awareness programs that is non-punitive and not directly involve the institution computer security staff. We feel that the trust between the security staff and the end user is very important and we want to avoid compromising the relationship. Given the student run IA program, we elected to organize the event at the staff level; however, it is implemented by the student chain of command. The student ISO's provided any corrective training immediately at and forwarded only statistical analysis to the Academy leadership. This no-threat, local resolution policy is a key aspect of the exercise. As described by David Jevans, chairman of the Anti-Phishing Working Group in Bank (2005), exercises like this might develop into a lack of trust of company emails.

The third consideration is to involve all stake holders. In our environment this included the Academy officer leadership and the students responsible to implement the security awareness and training program. We decided that to meet the objective of our first consideration, the student group involved would implement the exercise with faculty over-sight on the design. An unexpected benefit of exercising the real security response team arose from initiating the exercise at the student level. No information about the exercise initiation was passed on to the campus computer security team. When incidents of large amounts of phishing started to be reported, the staff had to execute just as if it were initiated externally. The Academy leadership saw this as a good opportunity to review and evaluate academy responses.

A fourth consideration that should be included regardless of the agency is involvement of the legal staff. There are two important facets that legal can provide opinions and guidance on; human subject research requirements and assessment on personal privacy.

The final general consideration is to construct the exercise to coexist with the organizational structure. We want the emails to be believable, but not lose their validity due to the nature of the organization. For example given the military nature of USMA, our students operate under a strongly enforced military hierarchy; however, they must be able to discern the difference between an appropriate and inappropriate request.

2.2. *Specific considerations*

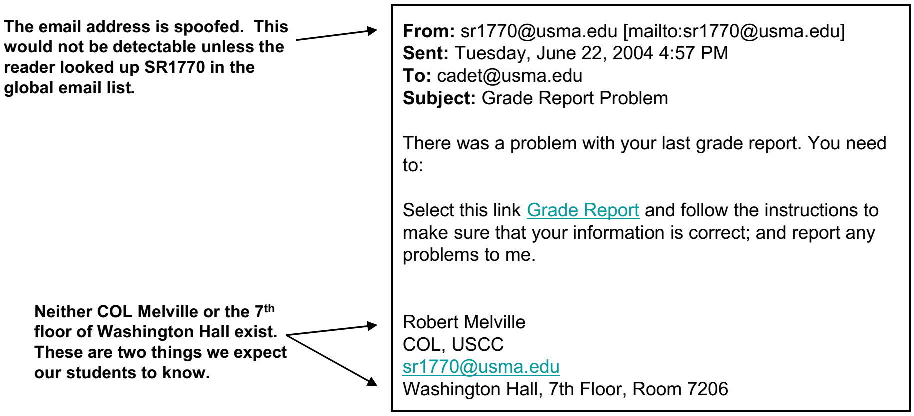
Once the general considerations are accounted for in the exercise design, we determined the specific requirements to make the exercise successful. Our first consideration was devising an email that a student would definitely be interested in opening, reading, and complying with. The objective is to provide the students with an email that they are enticed to open, however, if they read it carefully, they would realize it is not valid. One of our design decisions was to avoid common phishing email contents such as financial or services emails. We decided for our final email prototype that the student would be instructed to visit a website (hyperlink) to validate course grades. The second specific consideration, timing, fit nicely with the email content. For the prototype we constructed the email near the end of the semester when students would be most concerned about grade correctness. In general, the timing should be such that the test can accurately assess the defined objective. For example, if you are testing the long-term retention of security practices, the exercise should be conducted after a ''cool down'' period. Further if you are evaluating the increase in awareness over a time period, the exercise should be conducted at similar times for each implementation. The email in Fig. 1 is an example of the prototype.

The third specific consideration focused on the target. Prior to the exercise, a target population needs to be identified. For our prototype, a very small sample population across all classes was selected. The prototype (Carronade I) was limited in nature. The email referenced above was sent to 512 students at USMA (roughly 12% of the student body). Lastly, our fifth consideration, post event notification and follow-up mechanism, needed to be defined. Examples include, no notification, delayed notification to all ''targets'' (that may or may not include the rest of the student body), or immediate notification. In the prototype, immediate notification via an automated email was selected. In the next section we will discuss the evolution and addition of other email formats.

2.3. *Carronade evolution*

We conducted an assessment of the prototype and determined that the exercise could provide useful insights into the efficacy of our awareness and training program. Our focus for Carronade II was to develop a repeatable exercise that over time would serve as a yardstick to measure our programs. To that end, the exercise was conducted in September 2004. The timing of the exercise was set to ensure the new freshman class had an opportunity to become familiar with their computer and the email system at USMA. The exercise was conducted prior to any regularly scheduled IA training.

To validate that the email content was consistent with current training foci, we decided to seek input using a survey of information technology instructors. (This goes back to



The email address is spoofed. This would not be detectable unless the reader looked up SR1770 in the global email list.

From: sr1770@usma.edu [mailto:sr1770@usma.edu]
Sent: Tuesday, June 22, 2004 4:57 PM
To: cadet@usma.edu
Subject: Grade Report Problem

There was a problem with your last grade report. You need to:

Select this link Grade Report and follow the instructions to make sure that your information is correct; and report any problems to me.

Robert Melville
COL, USCC
sr1770@usma.edu
Washington Hall, 7th Floor, Room 7206

Neither COL Melville or the 7th floor of Washington Hall exist. These are two things we expect our students to know.

Fig. 1 – Example phishing email.

including the relevant stakeholders.) The survey asked the instructors to identify the top information assurance-related negative behaviors of their students. Using the survey, we developed the following requirements for Carronade II, incorporating the following requirements and four different styles of emails:

- The system must be able to deliver an email to all students.
- Each of the emails should be questionable enough to raise the suspicions of the end user.
- None of the emails, if opened outside of the USMA domain, would collect, track, or transmit any user data.
- The first email type asks the student to click on an embedded link in an HTML-encoded questionable email to fix a problem with a fictitious grade report; clicking on the link records information readily available from the browser.
- The second email type is identical to the first email except that the email asks the student to open the corresponding .html attachment.
- The third email type asks the student to click on a link that takes them to a web form that asks for sensitive information (i.e., their social security number).
- The fourth email type asks the students to click on a link, download an application and run the application. (This was implemented, however, in each exercise technical difficulties arising from cross domain issues prevented the success.)

As noted, a significant difference in Carronade II was the number of cadets targeted with the emails. This time the scope increased to the entire student body (minus those involved in the planning and execution). Of the total number of 4155 in the student body, 4118 received the emails. The email breakdown by type was: 1010 embedded link emails, 1014 attachment emails, 999 sensitive information emails, and 1095 download emails were sent out.

Carronade III was implemented functionally with very few changes from Carronade II. A similar sample population was used (everyone minus the exercise coordinators), however, the timing was changed to more closely follow the required IA training that each student receives. The date was selected to assess whether the recently received training produced lower "violations". The exercise was implemented in November 2005.

We chose a very similar email package – embedded link, attachment, sensitive information, and download emails. Unfortunately, different but similar problems plagued the download email and the results from this set are not included in the final analysis. The total emails sent 4136; were broken down as follows: 1006 embedded link, 1013 attachment emails, 1058 sensitive information emails, and 1059 download emails.

3. Technical implementation

We developed a generic design that is familiar to many web applications. The web application makes use of the Model-View-Controller design pattern separating data, display, and direction concerns within the program (Buschmann et al., 1996; Gamma et al., 1995). For our purposes we decided to use entirely open-source products to ensure that our work could be reproduced, as is, in any educational setting.

3.1. *Support architecture*

We chose Tomcat from Apache as the Web App Container which serves up both static HTML pages and dynamic Java Server Pages (JSP) (Apache Tomcat Java Servlet Container). The email server uses Apache James, an open-source email server solution (Apache James Email Server). We decided to use an object-relational mapping solution in our project because the database schema and models were changing often and we wanted to mitigate some of the time involved in maintaining the code base (Nock, 2004). We chose Hibernate which seems to be a fairly popular and successful ORM implementation (Hibernate Object-Relational Mapping Software). For our database server, we chose, MySQL. The database server and web app container were on the same physical computer (Dual processor, 3.2 GHz, 4 GB RAM, Win2K3). Our email server was on a separate computer, but could have been co-located with the database server and the web app container. Fig. 2 shows the object diagram for the technical implementation.

3.2. *Phishing email construction*

Our initial models were simple and evolved over time to represent the idea of a person which was used only for system administration reasons and the idea of a user's information. The student body is broken down into four equal sub-organizations called regiments. Each regiment is further broken down into eight companies. Each company has a similar make-up across all regiments with about 25 students from each year group.

The information model contains fields that may or may not be filled in depending on the type of email the student received. The type of email was the scheme. Most of the attributes of the class were related to the information that could be retrieved from a web browser in this case the strings starting with IP and ending with mime types. The reported attribute stored whether or not the recipient of the email had reported the incident to one of the information security support personnel. The opened attribute records whether or not the recipient actually clicked on the link or opened the attachment. Each of these model conform to the Java Bean standard for web app with private class-level instance variables and public methods to get and set the values of the variables.

As we looked to design a more robust suite of emails (from the basic type we used in the prototype) we adopted the

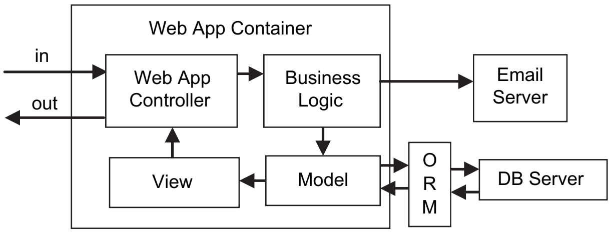


Fig. 2 – Web application logic.

recommendation described in Section 2.3. We selected four email phishing variants to employ; an embedded link, an attachment, a sensitive information request, and a download.

3.2.1. *Embedded link*

The first type of email which tried to get the user to click on an embedded link in an HTML-encoded email made use of a beacon, as shown in Fig. 1.

We define a beacon as a network link other than an anchor tag (i.e., <a></a> tags) in HTML over HTTP that when activated signals the existence and partial identity of a remote user. When the recipient selected the embedded link, their web browser (which is almost always Internet Explorer due to our homogenous software environment among students) opened and took them to an HTML page on the Web App Container. The link encoded their email address as a request parameter.

```
<ahref=‘http://hostname:8080/carronade/GradeReport.html?
Femail=“x00000@usma.edu”’>Grade Report</a>
```

The HTML page was a modified version of a standard Internet Explorer HTTP 404 page-not-found error. The only difference is that we added a script to get the email address from the request parameter, append each of the browser properties (which include OS, IP address, etc.) and append it to the end of an image link tag which did not go to an image, but to a Java Server Page which redirected directly to a servlet. The “^” and “|” characters were added to help aid in parsing the request parameter on the receiving servlet.

```
<script>
var text;
function submitInfo() {
   text = window.location.search + “&info = ”;
   for (var propertyName in navigator)
    text + = propertyName + “^” + navigator[propertyName] + “|”;
    document.images[0].src = document.images[0].src + text;
   return true;
   }
</script>
<body bgcolor=“white” onLoad=“submitInfo();”>
<img src=“http://hostname:8080/carronade/embeddedLink.jsp” />
```

In the servlet, the web app recorded each of the browser properties to include the recipient's email address and the fact that they had clicked on the link. The only barely noticeable indication that something was wrong on the recipient's end was that the image was never returned and results in a “red X”; however, this could be overcome by the servlet returning the correct image.

3.2.2. *Attachment*

The second type of email was the HTML attachment to an email. The email encouraged the recipient to open the attachment to fix some non-descript problem with their grades. The end result was exactly the same as the first type of email. Since the attachment was generated programmatically, we decided to keep it very small and sent the following one line of code in the HTML attachment.

```
<meta http-equiv=‘refresh’
content = ‘0;   url = http://hostname:8080/carronade/GradeReport.
html?email = x00000@usma.edu’>
```

If the recipient opens the attachment, a web browser opens automatically and is redirected to the GradeReport.html file as described in the embedded link attachment above. An example of the attachment email is shown in Fig. 1.

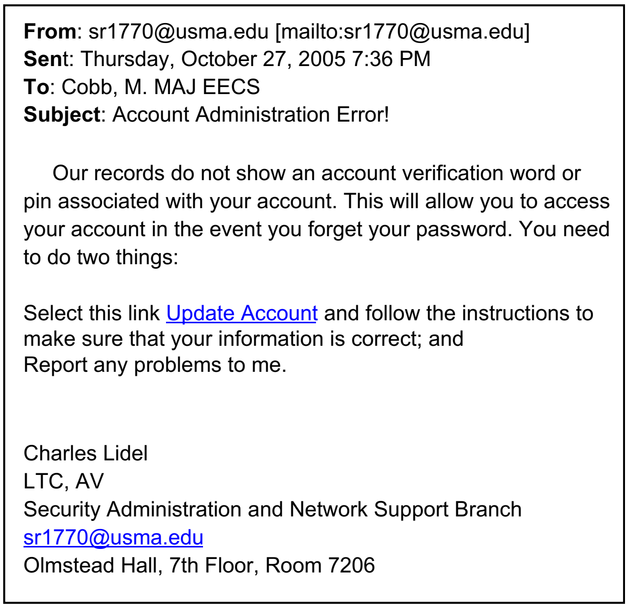
3.2.3. *Sensitive information*

The third type of email (Fig. 3) tried to social engineer the student into submitting their social security number for questionable reasons over an insecure connection to a questionable website. The objective in this scenario was not to record whether the student clicked a malicious link, but if they entered sensitive information.

It should be noted that we did not actually store any sensitive data (social security numbers we used for this exercise). In fact, the HTML code did not even send the social security numbers to the servlet as indicated by the absence of the name attribute in the second input tag.

```
<form action=“ssnServlet” method=“post”>
   User Id:<input type=“text” name=“userId” /><br/>
   Social Security Number:<input type=“text” /><br/>
   <input type=“submit” value=“Send Social Security” />
</form>
```

At the servlet, this was only registered if the user correctly typed in their user id which matches their email address. Because of this privacy concern the results missed tracking many of false negatives. There is also the possibility that a student might enter another student's user id which would lead to a false positive. Fortunately, the students who caught on to the fact that this was a bogus request typically entered in garbage in lieu of a friend's user id.



From: sr1770@usma.edu [mailto:sr1770@usma.edu]
Sent: Thursday, October 27, 2005 7:36 PM
To: Cobb, M. MAJ EECS
Subject: Account Administration Error!

Our records do not show an account verification word or pin associated with your account. This will allow you to access your account in the event you forget your password. You need to do two things:

Select this link Update Account and follow the instructions to make sure that your information is correct; and
Report any problems to me.

Charles Lidel
LTC, AV
Security Administration and Network Support Branch
sr1770@usma.edu
Olmstead Hall, 7th Floor, Room 7206

Fig. 3 – Sensitive information phishing email.

The fourth email type attempted to implement an email where the student was requested to download an application. Attempts were made in both Carronade II and III to implement this scenario, however, in each exercise technical difficulties arising from cross domain issues prevented the success.

4. Results

We elected to examine the results of the exercises in three facets. First, by overall percentage, by year, of the number of students that succumbed to the phishing emails. Second, we look at the distribution of failures by class for each exercise. Then finally, we look at the performance of a specific class over the two years.

4.1. *Failure percentage by exercise version*

As seen in Fig. 4, the failure rate in the prototype is very high. This is more than likely explained by the very small sample size (512). The results for 2004 and 2005 suggest minimal impact due to the recently conducted training. This, however, will require further data points to accurately draw this conclusion. It should be noted that in post exercise discussion with the students, the majority that did "fail" said they found the emails odd, however, responded anyway. The results indicate an 80% failure rate on the first prototype and approximately a 40% failure rate on the two subsequent exercises.

4.2. *Distribution by email type*

The breakout of failures (averaged over all Carronade interactions) by email type was interesting and proved to counter intuitive assumptions. The students were much more likely to open an attachment than they were to follow a link to a website or provide sensitive information. Fig. 5 shows that the failure rate was 38%, 50%, and 46% respectively. While the results do not indicate a significant propensity in any one area, they do provide a means to shape the discussion during email awareness training.

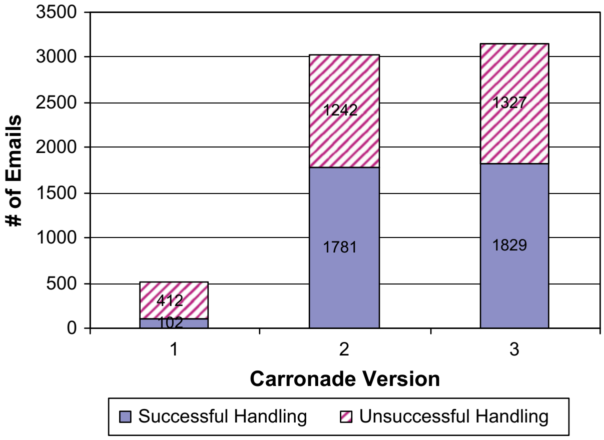


Fig. 4 – Success/failure rate by Carronade version.

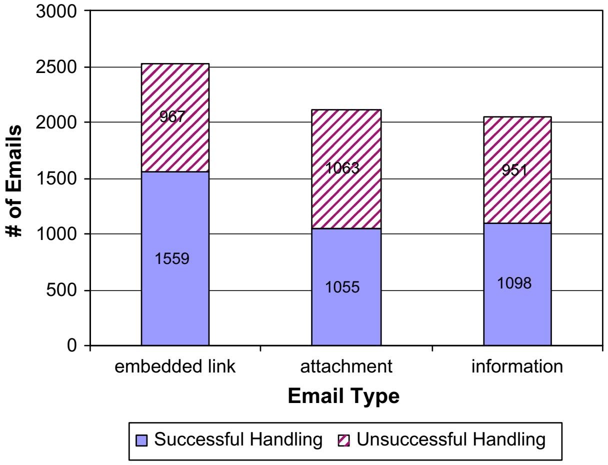


Fig. 5 – Email success/failure by email type.

4.3. *Distribution by class*

The analysis of performance of each class as a whole provides some insight into how effect over time the training is. In the first full Carronade, each class presented similar susceptibility to phishing attacks, as shown in Fig. 6.

The two outcomes we hoped to see was a reduction in the propensity of a student to fall victim to a phishing attack and also an increase in reporting that the phishing attack occurred. In the training we stress that phishing is going to happen and that what is to be reported is phishing attacks that appear to be targeted toward the USMA population.

Given our first goal of seeing a reduction in falling victim, we should expect to see the rate of replying to the phishing attacks decrease the longer a student is at USMA (receiving the training). As shown in Fig. 7, the rate of failure shows the declining function we hoped for.

A more detailed view of the failure rate as broken down by class year and exercise iteration show the improvement by class as a function of time is shown in Fig. 8. The bars indicate by shading a given class. The scale label indicates which Carronade iteration (1, 2, or 3) the data is from. As shown, the results indicated a better performance on each subsequent exercise for each class year.

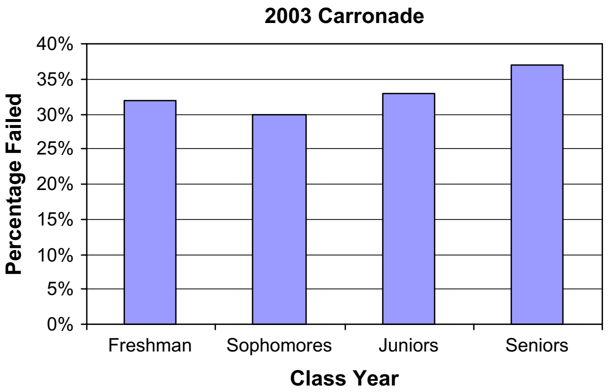


Fig. 6 – 2003 Carronade flat function.

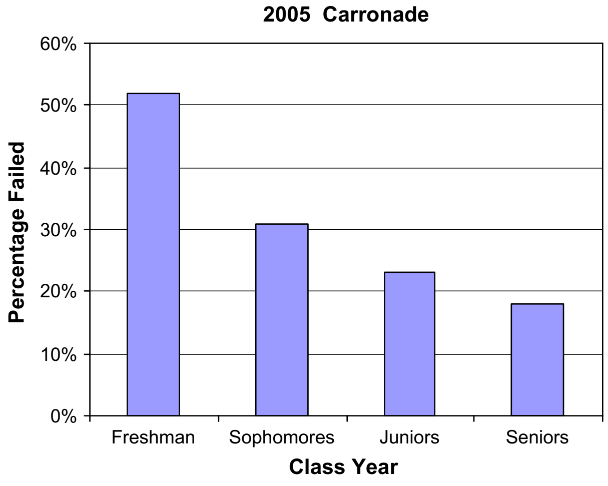


Fig. 7 – Class failure rate in 2005.

The second outcome was that the students would recognize a phishing attack that was specific to USMA, indicating a significant threat that warranted increased attention by the USMA staff. The initial statistics indicate that the longer a student is at USMA, the more likely they are to report phishing attempts. These results are shown in Fig. 9. Note that no specific questioning was done to determine if phishing attacks that are sent to the Internet population at large are also reported. However, anecdotal reports indicate that only USMA specific attacks are reported.

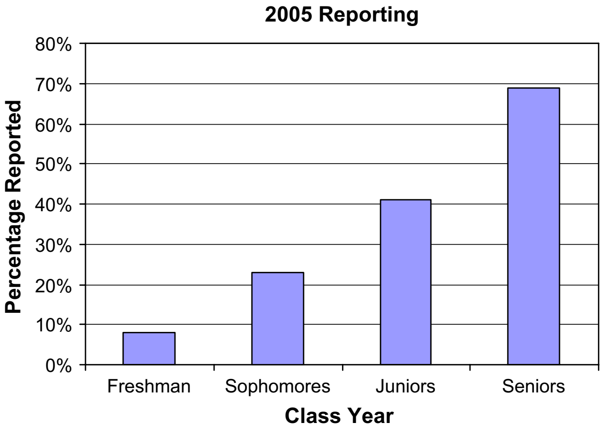


Fig. 9 – 2005 Phishing attack reporting.

5. Conclusions and future work

Our students continue to disclose information that should not be disclosed to an unauthorized user and expose themselves to malicious code by opening attachments. For the United States Military, this is important given the future requirement for operational security once the students graduate and enter the Army. This information will help us not only modify the IA awareness program, but also provide input to the other areas where operational security is important. The results of the analysis, however, do provide promising feedback that the IA awareness programs in place are having an impact on our students security consciousness.

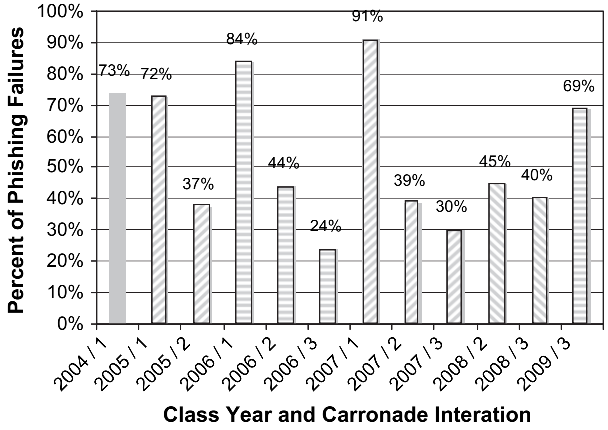


Fig. 8 – Results broken down by class/iteration.

The phishing exercises served to provide an insight into the awareness levels of our students and help us better focus our IA and awareness training. The results are still very immature; however, they provide an opportunity to look at the effectiveness of our programs. One might look at the assessment of the programs using an exercise as ‘‘poisoning the well’’, given the very fact that the exercises themselves may raise awareness, making it difficult to separate out any increased awareness due solely to the annual training. While this is true, when looking at the exercise from a bottom line – if our user’s awareness is increased, providing enhanced network security whatever the cause is a worthwhile cause.

We intend to continue the phishing email exercises, increasing the frequency to once every semester. One exercise will follow closely existing training; the second will be used to assess longer term retention. In addition to formalizing the scheduling of the exercise, we will expand the exercise to also include instant message (IM) traffic. The IM exercise will not occur in conjunction with the email phishing exercise.

REFERENCES

Anti-Phishing Working Group. July 2006 report, http://www.antiphishing.org/reports/apwg_report_july_2006.pdf [Accessed 13 September 2006].

Apache James Email Server. http://james.apache.org/.

Apache Tomcat Java Servlet Container. http://jakarta.apache.org/tomcat/.

Bank D. ‘Spear phishing’ tests educate people about online scams. Wall Street Journal August 2005;17.

Buschmann F, Meunier R, Rohnert H, Sommerlad P, Stal M. Pattern-oriented software architecture: a system of patterns. New York: Wiley; 1996. p. 125–43.

Dodge R, Hoffman L, Rosenberg T, Ragsdale D. Exploring a national cyber security exercise for universities. IEEE Security and Privacy 2005;September/October:52–8.

Dodge R, Ragsdale DJ, Reynolds C. Organization and training of a cyber security team. In: 2003 IEEE international conference

on systems, man and cybernetics, vol. 5; October 2003, p. 4306–11.
Gamma E, Helm R, Johnson R, Vlissides J. Design patterns. Addison-Wesley; 1995. p. 4–6.
Hibernate Object-Relational Mapping Software. http://www.hibernate.org/.
MySQL. http://www.mysql.com/.
Nock Clifton. Data access patterns. Addison-Wesley; 2004. p. 53–74.
Schepens W, James J. Architecture of a cyber defense competition. Systems, Man and Cybernetics October 2003;5:4300–5.
U.S. News, http://www.usnews.com/usnews/edu/elearning/lists/unwired_list.htm [accessed 13 September 2006].
Vigna G. Teaching hands-on network security: testbeds and live exercises. Journal of Information Warfare 2003;3(2):8–25.

Lt. Col. Dodge has served for over 19 years as an Aviation officer and is a member of the Army Acquisition Corps in the United States Army. His military assignments range from duties in an attack helicopter battalion during *Operation Just Cause* in the Republic of Panama to the United States Military Academy. Currently he is an Associate Professor permanently stationed at the United States Military Academy and the Director of the Information Technology and Operations Center (ITOC). Ron received his Ph.D. from George Mason University, Fairfax, Virginia in Computer Science, is a member of the ACM, IEEE, UPE, and IFIP, and is a CISSP. His current research focuses are Information Warfare, Network Deception, Security Protocols, Internet Technologies, and Performance Planning and Capacity Management. He is a frequent speaker at national and international IA conferences and has published many papers and articles on information assurance topics.

Lt. Col. Curtis A. Carver is an Army officer and Academy Professor at the United States Military Academy with over 20 years of service. He has served in a number of leadership positions including platoon leader, company commander, battalion operations officer, and division deputy G-6. His military awards include the Army Meritorious Service Medal with three oak leaf clusters, the Army Commendation with three oak leaf clusters, the Army Achievement Medal with three oak leaf clusters, and National Service Medal with star device.

Curt holds a Ph.D. in computer science and is a member of the ACM, IEEE, UPE, and PKP. He has over 90 academic works and a researcher in information assurance, adaptive hypermedia and computer science education. He has been a keynote speaker at several conferences including, most recently, the 2003 National Collegiate Engineering Symposium. Curt won the 1995 EDSIG Best Overall Paper Award, 1996 Frontiers in Education Ben Dasher Best Paper Award, 1996 and 1997 EDMEDIA Outstanding Paper Award, 1997 AFCIA Best Paper Award, and EISTA 2003 Best Track Paper Award and an honorable mentions at CCSC 2001. He is the Program Chair of the FISSEA conference.

Curt has also led numerous software engineering projects within the Army and academia including development of the TACWEB, DPASS, and Adaptive Hypermedia CS383 systems.

Dr. Aaron J. Ferguson is currently a Program Manager in the Advanced Network Operations office at the National Security Agency. Prior to his current assignment, Dr. Ferguson was the National Security Agency Fellow in the Department of Electrical Engineering and Computer Science at the United States Military Academy at West Point where he taught Software System Design and other security-related courses. Dr. Ferguson is an Electrical Engineering graduate of Howard University where he focused on control systems. He received an M.S. in Operations Research from the University of New Haven where he focused on software reliability modeling. He earned an M.S. and Ph.D. in Applied Statistics from the University of Delaware where he focused on system simulation and modeling. Dr. Ferguson is a Certified Information System Security Professional (CISSP) and his research interests include Insider Threat Detection and Modeling, Security Awareness, Training, and Education (SATE), XML Security, Cross- Security Domain Information Exchange Architectures. Dr. Ferguson has taught Information Assurance course overseas in private industry and as a Senior Consultant for PricewaterhouseCoopers, LLC, has provided risk management consultancy to several Fortune 500 companies.

Computers & Security

A privacy-preserving clustering approach toward secure and effective data analysis for business collaboration[☆]

Stanley R.M. Oliveira[a,*], *Osmar R. Zaïane*[b]

[a]*Embrapa Informática Agropecuária, Av. André Tosello, 209, 13083-886 Campinas, SP, Brasil*
[b]*Department of Computing Science, University of Alberta, Edmonton, AB, Canada T6G 2E8*

ARTICLE INFO

Article history:
Received 22 July 2005
Accepted 3 August 2006

Keywords:
Privacy-preserving data mining
Privacy-preserving clustering
Dimensionality reduction
Random projection
Privacy-preserving clustering over centralized data
Privacy-preserving clustering over vertically partitioned data

ABSTRACT

The sharing of data has been proven beneficial in data mining applications. However, privacy regulations and other privacy concerns may prevent data owners from sharing information for data analysis. To resolve this challenging problem, data owners must design a solution that meets privacy requirements and guarantees valid data clustering results. To achieve this dual goal, we introduce a new method for privacy-preserving clustering called Dimensionality Reduction-Based Transformation (DRBT). This method relies on the intuition behind random projection to protect the underlying attribute values subjected to cluster analysis. The major features of this method are: (a) it is independent of distance-based clustering algorithms; (b) it has a sound mathematical foundation; and (c) it does not require CPU-intensive operations. We show analytically and empirically that transforming a data set using DRBT, a data owner can achieve privacy preservation and get accurate clustering with a little overhead of communication cost.

1. Introduction

In the business world, data clustering has been used extensively to find the optimal customer targets, improve profitability, market more effectively, and maximize return on investment supporting business collaboration (Lo, 2002; Berry and Linoff, 1997). Often combining different data sources provides better clustering analysis opportunities. For example, it does not suffice to cluster customers based on their purchasing history, but combining purchasing history, vital statistics and other demographic and financial information for clustering purposes can lead to better and more accurate customer behaviour analysis. However, this means sharing data between parties.

Despite its benefits to support both modern business and social goals, clustering can also, in the absence of adequate safeguards, jeopardize individuals' privacy. The fundamental question addressed in this paper is: *how can data owners protect personal data shared for cluster analysis and meet their needs to support decision making or to promote social benefits?* To address this problem, data owners must not only meet privacy requirements but also guarantee valid clustering results.

Clearly, achieving privacy preservation when sharing data for clustering poses new challenges for novel uses of data

[☆] A preliminary version of this work appeared in the Workshop on Privacy and Security Aspects of Data Mining in conjunction with ICDM 2004, Brighton, UK, November 2004.

[*] *Corresponding author*. Embrapa Informática Agropecuária, Av. André Tosello, 209, 13083-886 Campinas, SP, Brasil.
E-mail addresses: stanley@cnptia.embrapa.br (S.R.M. Oliveira), zaiane@cs.ualberta.ca (O.R. Zaïane).

mining technology. Each application poses a new set of challenges. Let us consider two real-life examples in which the sharing of data poses different constraints:

- Two organizations, an Internet marketing company and an on-line retail company, have data sets with different attributes for a common set of individuals. These organizations decide to share their data for clustering to find the optimal customer targets so as to maximize return on investments. How can these organizations learn about their clusters using each other's data without learning anything about the attribute values of each other?
- Suppose that a hospital shares some data for research purposes (e.g., to group patients who have a similar disease); the hospital's security administrator may suppress some identifiers (e.g., name, address, phone number, etc.) from patient records to meet privacy requirements. However, the released data may not be fully protected. A patient record may contain other information that can be linked with other data sets to re-identify individuals or entities (Samarati, 2001; Sweeney, 2002). How can we identify groups of patients with a similar pathology or characteristics without revealing the values of the attributes associated with them?

The above scenarios describe two different problems of privacy-preserving clustering (PPC). We refer to the former as *PPC over centralized data* and the latter as *PPC over vertically partitioned data*. To address these scenarios, we introduce a new PPC method called Dimensionality Reduction-Based Transformation (DRBT). This method allows data owners to find a trade-off between privacy, accuracy, and communication cost. Communication cost is the cost (typically in size) of the data exchanged between parties in order to achieve secure clustering.

Dimensionality reduction techniques have been studied in the context of pattern recognition (Fukunaga, 1990), information retrieval (Bingham and Mannila, 2001; Faloutsos and Lin, 1995; Jagadish, 1991), and data mining (Fern and Brodley, 2003; Faloutsos and Lin, 1995). To the best of our knowledge, dimensionality reduction has not been used in the context of data privacy in any detail, except in Oliveira and Zaïane (2004).

Although there exists a number of methods for reducing the dimensionality of data, such as feature extraction methods (Kaski, 1999), multi-dimensional scaling (Young, 1987) and principal component analysis (PCA) (Fukunaga, 1990), this paper focuses on random projection, a powerful method for dimensionality reduction. The accuracy obtained after the dimensionality has been reduced, using random projection, is almost as good as the original accuracy (Kaski, 1999; Achlioptas, 2001; Bingham and Mannila, 2001). More formally, when a vector in d-dimensional space is projected onto a random k-dimensional subspace, the distances between any pair of points are not distorted by more than a factor of $(1 \pm \epsilon)$, for any $0 < \epsilon < 1$, with probability $O(1/n^2)$, where n is the number of objects under analysis (Johnson and Lindenstrauss, 1984).

The motivation for exploring random projection is based on the following aspects. First, it is a general data reduction technique. In contrast to the other methods, such as PCA, random projection does not use any defined interestingness criterion to optimize the projection. Second, random projection has shown to have promising theoretical properties for high-dimensional data clustering (Fern and Brodley, 2003; Bingham and Mannila, 2001). Third, despite its computational simplicity, random projection does not introduce a significant distortion in the data. Finally, the dimensions found by random projection are not a subset of the original dimensions but rather a transformation, which is relevant for privacy preservation.

In this work, random projection is used to mask the underlying attribute values subjected to clustering, protecting them from being revealed. In tandem with the benefit of privacy preservation, the method DRBT benefits from the fact that random projection preserves the distances (or similarities) between data objects quite nicely, which is desirable in cluster analysis. We show analytically and experimentally that using DRBT, a data owner can meet privacy requirements without losing the benefit of clustering.

The major features of our method, DRBT, are: (a) it is independent of distance-based clustering algorithms; (b) it has a sound mathematical foundation; and (c) it does not require CPU-intensive operations.

This paper is organized as follows. In Section 2, we provide the basic concepts that are necessary to understand the issues addressed in this paper. In Section 3, we describe the research problem employed in our study. In Section 4, we introduce our method DRBT to address PPC over centralized data and over vertically partitioned data. The experimental results are presented in Section 5. Related work is reviewed in Section 6. Finally, Section 7 presents our conclusions.

2. Background

In this section, we briefly review the basic concepts that are necessary to understand the issues addressed in this paper.

2.1. *Data matrix*

Objects (e.g., individuals, observations, events) are usually represented as points (vectors) in a multi-dimensional space. Each dimension represents a distinct attribute describing the object. Thus, objects are represented as an $m \times n$ matrix D, where there are m rows, one for each object, and n columns, one for each attribute. This matrix may contain binary, categorical, or numerical attributes. It is referred to as a data matrix, as can be seen in Fig. 1.

The attributes in a data matrix are sometimes transformed before being used. The main reason is that different attributes may be measured on different scales (e.g., centimeters and kilograms). When the range of values differs widely from attribute to attribute, attributes with large range can influence the results of the cluster analysis. For this reason, it is common to

$$\mathrm{D} = \begin{bmatrix} a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & \dots & a_{2k} & \dots & a_{2n} \\ \vdots & & \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mk} & \dots & a_{mn} \end{bmatrix}$$

Fig. 1 – The data matrix structure.

standardize the data so that all attributes are on the same scale. There are many methods for data normalization (Han et al., 2001). We review only two of them in this section: *min–max normalization* and *z-score normalization*.

Min–max normalization performs a linear transformation on the original data. Each attribute is normalized by scaling its values so that they fall within a specific range, such as 0.0 and 1.0.

When the actual minimum and maximum of an attribute are unknown, or when there are outliers that dominate the min–max normalization, z-score normalization (also called zero-mean normalization) should be used. In this case, the normalization is performed by subtracting the mean from each attribute value and then dividing the result by the standard deviation of this attribute.

2.2. *Dissimilarity matrix*

A dissimilarity matrix stores a collection of proximities that are available for all pairs of objects. This matrix is often represented by an $m \times m$ table. In Fig. 2, we can see the dissimilarity matrix D_M corresponding to the data matrix D in Fig. 1, where each element $d(i, j)$ represents the difference or dissimilarity between objects i and j.

In general, $d(i, j)$ is a non-negative number that is close to zero when the objects i and j are very similar to each other, and becomes larger the more they differ.

Several distance measures could be used to calculate the dissimilarity matrix of a set of points in d-dimensional space (Han et al., 2001). The Euclidean distance is the most popular distance measure. If $i = (x_{i1}, x_{i2}, \ldots, x_{in})$ and $j = (x_{j1}, x_{j2}, \ldots, x_{jn})$ are n-dimensional data objects, the Euclidean distance between i and j is given by:

$$d(i,j) = \left[\sum_{k=1}^{n} \left| x_{ik} - x_{jk} \right|^2 \right]^{\frac{1}{2}} \tag{1}$$

The Euclidean distance satisfies the following constraints:

- $d(i, j) \geq 0$: distance is a non-negative number.
- $d(i, i) = 0$: the distance of an object to itself.
- $d(i, j) = d(j, i)$: distance is a symmetric function.
- $d(i, j) \leq d(i, k) + d(k, j)$: distance satisfies the triangular inequality.

2.3. *Random projection*

In many applications of data mining, the high dimensionality of the data restricts the choice of data processing methods. Examples of such applications include market basket data, text classification, and clustering. In these cases, the dimensionality is large due to either a wealth of alternative products, a large vocabulary, or an excessive number of attributes to be analyzed in Euclidean space, respectively.

$$D_M = \begin{bmatrix} 0 & & & & \\ d(2,1) & 0 & & & \\ d(3,1) & d(3,2) & 0 & & \\ \ldots & \ldots & \ldots & & \\ d(m,1) & d(m,2) & \ldots & \ldots & 0 \end{bmatrix}$$

Fig. 2 – The dissimilarity matrix structure.

When data vectors are defined in a high-dimensional space, it is computationally intractable to use data analysis or pattern recognition algorithms that repeatedly compute similarities or distances in the original data space. It is therefore necessary to reduce the dimensionality before, for instance, clustering the data (Kaski, 1999; Fern and Brodley, 2003).

The goal of the methods designed for dimensionality reduction is to map d-dimensional objects into k-dimensional objects, where $k \ll d$ (Kruskal and Wish, 1978). These methods map each object to a point in a k-dimensional space minimizing the stress function:

$$\text{stress}^2 = \frac{\sum_{i,j} \left(\hat{d}_{ij} - d_{ij} \right)^2}{\sum_{i,j} d_{ij}^2} \tag{2}$$

where d_{ij} is the dissimilarity measure between objects i and j in a d-dimensional space, and $\hat{d}_{ij}$ is the dissimilarity measure between objects i and j in a k-dimensional space. The function *stress* gives the relative error that the distances in k–d space suffer from, on the average.

One of the methods designed for dimensionality reduction is random projection. This method has been shown to have promising theoretical properties since the accuracy obtained after the dimensionality has been reduced, using random projection, is almost as good as the original accuracy. Most importantly, the rank order of the distances between data points is meaningful (Kaski, 1999; Achlioptas, 2001; Bingham and Mannila, 2001). The key idea of random projection arises from the Johnson and Lindenstrauss (1984) lemma: "if points in a vector space are projected onto a randomly selected subspace of suitably high dimension, then the distances between the points are approximately preserved."

Lemma 1. *Given $\epsilon > 0$ and an integer n, let k be a positive integer such that $k \geq k_0 = O(\epsilon^{-2} \log n)$. For every set P of n points in R^d there exists $f : R^d \to R^k$ such that for all $u, v \in P$*

$$(1 - \epsilon)\|u - v\|^2 \leq \|f(u) - f(v)\|^2 \leq (1 + \epsilon)\|u - v\|^2.$$

The classic result of Johnson and Lindenstrauss (1984) asserts that any set of n points in d-dimensional Euclidean space can be embedded into k-dimensional space, where k is logarithmic in n and independent of d. Thus to get the most of random projection, the following constraint must be satisfied: $k \geq k_0 = O(\epsilon^{-2} \log n)$.

A random projection from d dimensions to k dimensions is a linear transformation represented by a $d \times k$ matrix R, which is generated by first setting each entry of the matrix to a value drawn from an i.i.d. $\sim N(0, 1)$ distribution (i.e., zero mean and unit variance) and then normalizing the columns to unit length. Given a d-dimensional data set represented as an $n \times d$ matrix D, the mapping $D \times R$ results in a reduced-dimension data set D', i.e.,

$$D'_{n \times k} = D_{n \times d} R_{d \times k} \tag{3}$$

Random projection is computationally very simple. Given the random matrix R and projecting the $n \times d$ matrix D into k

dimensions is of the order O(*ndk*), and if the matrix *D* is sparse with about *c* nonzero entries per column, the complexity is of the order O(*cnk*) (Papadimitriou et al., 1998).

After applying random projection to a data set, the distance between two *d*-dimensional vectors *i* and *j* is approximated by the scaled Euclidean distance of these vectors in the reduced space as follows:

$$\sqrt{d/k}\|R_i - R_j\| \tag{4}$$

where *d* is the original and *k* the reduced dimensionality of the data set. The scaling term $\sqrt{d/k}$ takes into account the decrease in the dimensionality of the data. To satisfy Lemma 1, the random matrix *R* must hold the follow constraints:

- The columns of the random matrix *R* are composed of orthonormal vectors, i.e., they have unit length and are orthogonal.
- The elements r_{ij} of *R* have zero mean and unit variance.

Clearly, the choice of the random matrix *R* is one of the key points of interest. The elements r_{ij} of *R* are often Gaussian distributed, but this need not to be the case. Achlioptas (2001) showed that the Gaussian distribution can be replaced by a much simpler distribution, as follows:

$$r_{ij} = \sqrt{3} \times \begin{cases} +1 & \text{with probability } 1/6 \\ 0 & \text{with probability } 2/3 \\ -1 & \text{with probability } 1/6 \end{cases} \tag{5}$$

In fact, practically all zero mean, unit variance distributions of r_{ij} would give a mapping that still satisfies the Johnson–Lindenstrauss lemma. Achlioptas' result means further computational savings in database applications since the computations can be performed using integer arithmetic.

3. Privacy-preserving clustering: problem definition

The goal of privacy-preserving clustering is to protect the underlying attribute values of objects subjected to clustering analysis. In doing so, the privacy of individuals would be protected.

The problem of privacy preservation in clustering can be stated as follows: let *D* be a relational database and *C* a set of clusters generated from *D*. The goal is to transform *D* into *D*′ so that the following restrictions hold:

- A transformation $\mathcal{T}$ when applied to *D* must preserve the privacy of individual records, so that the released database *D*′ conceals the values of confidential attributes, such as salary, disease diagnosis, credit rating, and others.
- The similarity between objects in *D*′ must be the same as that one in *D*, or just slightly altered by the transformation process. Although the transformed database *D*′ looks very different from *D*, the clusters in *D* and *D*′ should be as close as possible since the distances between objects are preserved or marginally changed.

We will approach the problem of PPC by first dividing it into two sub-problems: PPC over centralized data and PPC over vertically partitioned data. In the centralized data approach, different entities are described with the same schema in a unique centralized data repository, while in a vertical partition, the attributes of the same entities are split across the partitions. We do not address the case of horizontally partitioned data.

3.1. *PPC over centralized data*

In this scenario, two parties, **A** and **B**, are involved, party **A** owning a data set *D* and party **B** wanting to mine it for clustering. In this context, the data are assumed to be a matrix $D_{m\times n}$, where each of the *m* rows represents an object, and each object contains values for each of the *n* attributes.

Before sharing the data set *D* with party **B**, party **A** must transform *D* to preserve the privacy of individual data records. After transformation, the attribute values of an object in *D* would look very different from the original. However, the transformation applied to *D* must not jeopardize the similarity between objects. Therefore, miners would rely on the transformed data to build valid results, i.e., clusters. Our second real-life motivating example, in Section 1, is a particular case of PPC over centralized data.

3.2. *PPC over vertically partitioned data*

Consider a scenario wherein *k* parties, such that $k \geq 2$, have different attributes for a common set of objects, as mentioned in the first real-life example, in Section 1. Here, the goal is to do a join over the *k* parties and cluster the common objects. The data matrix for this case is given as follows:

$$\begin{array}{cccc} \vdash \text{Party 1} \dashv & \vdash \text{Party 2} \dashv & & \vdash \text{Party } k \dashv \\ \end{array}$$

$$\begin{bmatrix} a_{11}\ldots a_{1i} & a_{1i+1}\ldots a_{1j} & \cdots & a_{1p+1}\ldots a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\ldots a_{mi} & a_{mi+1}\ldots a_{mj} & \cdots & a_{mp+1}\ldots a_{mn} \end{bmatrix} \tag{6}$$

Note that, after doing a join over the *k* parties, the problem of PPC over vertically partitioned data becomes a problem of PPC over centralized data. For simplicity, we do not consider communication cost here since this issue is addressed later.

In our model for PPC over vertically partitioned data, one of the parties is the central one which is in charge of merging the data and finding the clusters in the merged data. After finding the clusters, the central party would share the clustering results with the other parties. The challenge here is how to move the data from each party to a central party concealing the values of the attributes of each party. However, before

Table 1 – Thread of selecting the attributes on the party$_k$ side

Steps to select the attributes for clustering on the party$_k$ side
1. Negotiate the attributes for clustering before the sharing of data. 2. Wait for the list of attributes available in party$_c$. 3. Upon receiving the list of attributes from party$_c$: (a) Select the attributes of the objects to be shared.

Table 2 – Thread of selecting the objects on the party$_k$ side

Steps to select the list of objects on the party$_k$ side
1. Negotiate the list of *m* objects before the sharing of data. 2. Wait for the list of *m* object IDs. 3. Upon receiving the list of *m* object IDs from party$_c$: (a) Select *m* objects to be shared; (b) Transform the attribute values of the *m* objects; (c) Send the transformed *m* objects to party$_c$.

moving the data to a central party, each party must transform its data to protect the privacy of the attribute values. We assume that the existence of an object (ID) should be revealed for the purpose of the join operation, but the values of the associated attributes are private.

3.3. *The communication protocol*

To address the problem of PPC over vertically partitioned data, we need to design a communication protocol. This protocol is used between two parties: the first party is the central one and the other represents any of the $k-1$ parties, assuming that we have k parties. We refer to the central party as party$_c$ and any of the other parties as party$_k$. There are two threads on the party$_k$ side, one for selecting the attributes to be shared, as can be seen in Table 1, and the other for selecting the objects before the sharing of data, as can be seen in Table 2.

4. The Dimensionality Reduction-Based Transformation

In this section, we show that the triple-goal of achieving privacy preservation and valid clustering results at a reduced communication cost in PPC can be accomplished by random projection. By reducing the dimensionality of a data set to a sufficiently small value, one can find a trade-off between privacy, accuracy, and communication cost. We refer to this solution as the Dimensionality Reduction-Based Transformation (DRBT).

4.1. *General assumptions*

The solution to the problem of PPC draws the following assumptions:

- The data matrix D subjected to clustering contains only numerical attributes that must be transformed to protect individuals' data values before the data sharing for clustering occurs.
- In PPC over centralized data, the identity of an object (ID) must be replaced by a fictitious identifier. In PPC over vertically partitioned data, the IDs of the objects are used for the join purposes between the parties involved in the solution, and the existence of an object at a site is not considered private.

One interesting characteristic of the solution based on random projection is that, once the dimensionality of a database is reduced, the attribute names in the released database are irrelevant. We refer to the released database as a *disguised database*, which is shared for clustering.

4.2. *PPC over centralized data*

To address PPC over centralized data, the DRBT performs three major steps before sharing the data for clustering:

- *Step 1 – suppressing identifiers*: attributes that are not subjected to clustering (e.g., address, phone number, etc.) are suppressed.
- *Step 2 – reducing the dimension of the original data set*: after preprocessing the data according to *Step 1*, an original data set D is then transformed into the disguised data set D' using random projection.
- *Step 3 – computing the stress function*: this function is used to determine that the accuracy of the transformed data is marginally modified, which guarantees the usefulness of the data for clustering. A data owner can compute the stress function using Eq. (2).

To illustrate how this solution works, let us consider the sample relational database in Table 3. This sample contains real data from the Cardiac Arrhythmia Database available at the UCI Repository of Machine Learning Databases (Blake and Merz, 1998). The attributes for this example are: *age*, *weight*, *h_rate* (number of heart beats per minute), *int_def* (number of intrinsic deflections), *QRS* (average of QRS duration in milliseconds), and *PR_int* (average duration between onset of *P* and *Q* waves in milliseconds).

We are going to reduce the dimension of this data set from 6 to 3, one at a time, and compute the error (stress function). To reduce the dimension of this data set, we apply Eq. (3). In this example, the original data set corresponds to the matrix D. We compute a random matrix R_1 by setting each entry of the matrix to a value drawn from an independent and identically distributed (i.i.d.) $N(0, 1)$ distribution and then normalizing the columns to unit length. We also compute a random matrix R_2 where each element r_{ij} is computed using Eq. (5). We transform D into D' using both R_1 and R_2. The random transformation RP$_1$ refers to the random projection using R_1, and RP$_2$ refers to the random projection using R_2.

The relative error that the distances in 6–3 space suffer from, on the average, is computed using Eq. (2). Table 4 shows the values of the error using RP$_1$ and RP$_2$. In this table, k represents the number of dimensions in the disguised database D'.

In this case, we have reduced the dimension of D from 6 to 3, i.e., the transformed data set has only 50% of the dimensions in the original data set. Note that the error is relatively small for

Table 3 – A cardiac arrhythmia database

ID	Age	Weight	h_rate	int_def	QRS	PR_int
123	75	80	63	32	91	193
342	56	64	53	24	81	174
254	40	52	70	24	77	129
446	28	58	76	40	83	251
286	44	90	68	44	109	128

Table 4 – The relative error that the distances in 6–3 space suffer from

Transformation	$k=6$	$k=5$	$k=4$	$k=3$
RP_1	0.0000	0.0223	0.0490	0.2425
RP_2	0.0000	0.0281	0.0375	0.1120

both RP_1 and RP_2, especially for RP_2. However, this error is minimized when the random projection is applied to high-dimensional data sets, as can be seen in Fig. 4, in Section 5.4.

After applying random projection to a data set, the attribute values of the transformed data set are completely disguised to preserve the privacy of individuals. Table 5 shows the attribute values of the transformed database with three dimensions, using both RP_1 and RP_2. In this table, we have the attributes labeled Att1, Att2, and Att3 since we do not know the labels for the disguised data set. Using random projection, one cannot select the attributes to be reduced beforehand. The attributes are reduced randomly. More formally, $\forall i$ if $\text{Attr}_i \in D'$, then $\text{Attr}_i \notin D$.

As can be seen in Table 5, the attribute values are entirely different from those in Table 3.

4.3. *PPC over vertically partitioned data*

The solution for PPC over vertically partitioned data is a generalization of the solution for PPC over centralized data. In particular, if we have k parties involved in this case, each party must apply the random projection over its data set and then send the reduced data matrix to a central party. Note that any of the k parties can be the central one.

When k parties ($k \geq 2$) share some data for PPC over vertically partitioned data, these parties must satisfy the following constraints:

- *Agreement*: the k parties must follow the communication protocol described in Section 3.3.
- *Mutual exclusivity*: we assume that the attribute splits across the k parties are mutually exclusive. More formally, if $A(D_1)$, $A(D_2)$, …, $A(D_k)$ are a set of attributes of the k parties, $\forall i \neq j$ $A(D_i) \cap A(D_j) = \varnothing$. The only exception is that IDs are shared for the joining purpose.

The solution based on random projection for PPC over vertically partitioned data is performed as follows:

- *Step 1 – individual transformation*: if k parties, $k \geq 2$, share their data in a collaborative project for clustering, each party k_i must transform its data according to the steps in Section 4.2.
- *Step 2 – data exchanging or sharing*: once the data are disguised by using random projection, the k parties are able to exchange the data among themselves. However, one party could be the central one to aggregate and cluster the data.
- *Step 3 – sharing clustering results*: after the data have been aggregated and mined in a central party k_i, the results could be shared with the other parties.

Table 5 – Disguised data set D' using RP_1 and RP_2

ID	D' using RP_1			D' using RP_2		
	Att1	Att2	Att3	Att1	Att2	Att3
123	−50.40	17.33	12.31	−55.50	−95.26	−107.96
342	−37.08	6.27	12.22	−51.00	−84.29	−83.13
254	−55.86	20.69	−0.66	−65.50	−70.43	−66.97
446	−37.61	−31.66	−17.58	−85.50	−140.87	−72.74
286	−62.72	37.64	18.16	−88.50	−50.22	−102.76

4.4. *How secure is the DRBT?*

In the previous sections, we showed that transforming a database using random projection is a promising solution for PPC over centralized data and consequently for PPC over vertically partitioned data since the similarities between objects are marginally changed. Now we show that random projection also has promising theoretical properties for privacy preservation. In particular, we demonstrate that a random projection from d dimensions to k, where $k \ll d$, is a non-invertible transformation.

Lemma 2. *A random projection from d dimensions to k dimensions, where $k \ll d$, is a non-invertible linear transformation.*

Proof. A classic result from Linear Algebra asserts that there is no invertible linear transformation between Euclidean spaces of different dimensions (Auer, 1991). Thus, if there is an invertible linear transformations from R^m to R^n, then the constraint $m = n$ must hold. A random projection is a linear transformation from R^d to R^k, where $k \ll d$. Hence, a random projection from d dimensions to k dimensions is a non-invertible linear transformation. □

Even when sufficient care is taken, a solution that adheres to DRBT can be still vulnerable to disclosure. For instance, if an adversary knows the positions of $d+1$ points (where d is the number of dimensions), and the distances between these points, then one can make some estimates of the coordinates of all points. Caetano (2004) shows that if an adversary knows the dissimilarity matrix of a set of points and the coordinates of $d+1$ points, where d is the number of dimensions of the data points, it is possible to disclose the entire data set. However, this result holds if and only if the $d+1$ points do not lie in a $(d-1)$-dimensional vector subspace.

To illustrate Caetano's lemma, let us consider a particular case in R^2, as can be seen in Fig. 3. In this example, suppose that three point $(d+1)$ objects $(d=2)$ and their distances are known. If the center of the dashed circle does not lie in the same straight line, the $(d-1)$-dimensional vector subspace, defined by the centers of the other two circles, the intersection set has at most one point (the one pointed to by the arrow). Thus, if one adversary has the distances of other p points to these three points in Fig. 3, s/he can determine the coordinates of the p points.

It is important to note that the violation of the solution that adheres to DRBT becomes progressively harder as the number of attributes (dimensions) in a database increases since an adversary would need to know $d+1$ points to disclose the original data. On the other hand, when the number of dimensions grows, the accuracy regarding the distances between points is improved.

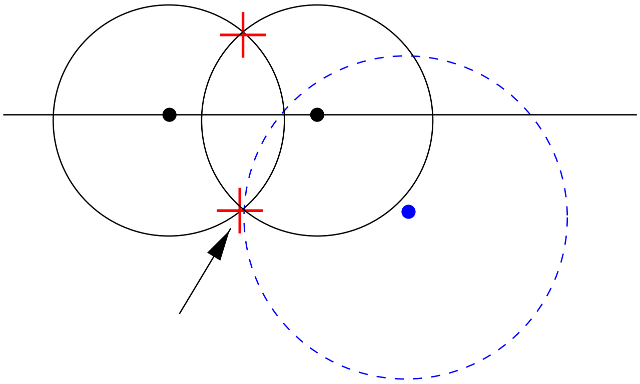


Fig. 3 – An example of Caetano's lemma in R^2 **(Caetano, 2004).**

4.5. *The complexity of the DRBT*

One of the major benefits of a solution that adheres to the DRBT is the communication cost to send a disguised data set from one party to a central one. In general, a disguised data matrix is of size $m \times k$, where m is the number of objects and k is the number of attributes (dimensions). The complexity of DRBT is of the order $O(m \times k)$, however, $k \ll m$.

To quantify the communication cost of one solution, we consider the number of bits or words required to transmit a data set from one party to a central or third party. Using DRBT, the bit communication cost to transmit a data set from one party to another is $O(mlk)$, where l represents the size (in bits) of one element of the $m \times k$ disguised data matrix.

5. Experimental results

In this section, we empirically validate our method DRBT. We start by describing the real data sets used in our experiments. We then describe the methodology and the evaluation approach used to validate our method. Subsequently, we study the effectiveness of our method to address PPC over centralized data and PPC over vertically partitioned data. We conclude this section discussing the main lessons learned from our experiments.

5.1. *Data sets*

We validated our method DRBT for privacy-preserving clustering using five real data sets. These data sets are described as follows:

1. *Accidents*: this data set concerning traffic accidents was obtained from the National Institute of Statistics (NIS) for the region of Flanders in Belgium. There are 340,183 traffic accident records included in the data set. We used 18 columns of this data set after removing the missing values.
2. *Mushroom*: this data set is available at the UCI Repository of Machine Learning Databases (Blake and Merz, 1998). Mushroom contains records drawn from The Audubon Society Field Guide to North American Mushrooms. There are 8124 records and 23 numerical attributes.
3. *Chess*: the format for instances in this database is a sequence of 37 attribute values. Each instance is a board-description of a chess endgame. The first 36 attributes describe the board. The last (37th) attribute is the classification: ''win'' or ''nowin''. Chess is available at the UCI Repository of Machine Learning Databases (Blake and Merz, 1998) and contains 3196 records. There is no missing value in this data set.
4. *Connect*: this database contains all legal 8-ply positions in the game of connect-4 in which neither player has won yet, and in which the next move is not forced. Connect is composed of 67,557 records and 43 attributes without missing values. This data set is also available at the UCI Repository of Machine Learning Databases (Blake and Merz, 1998).
5. *Pumsb*: the Pumsb data set contains census data for population and housing. This data set is available at http://www.almaden.ibm.com/software/quest. There are 49,046 records and 74 attribute values without missing values.

Table 6 shows the summary of the data sets used in our experiments. The columns represent, respectively, the database name, the total number of records, and the number of attributes in each data set.

Table 6 – A summary of the data sets used in our experiments

Dataset	No. of records	No. of attributes
Accidents	340,183	18
Mushroom	8124	23
Chess	3196	37
Connect	67,557	43
Pumsb	49,046	74

5.2. *Methodology*

We performed two series of experiments to evaluate the effectiveness of DRBT when addressing PPC over centralized data and PPC over vertically partitioned data. Our evaluation approach focused on the overall quality of generated clusters after dimensionality reduction. One question that we wanted to answer was:

> *What is the quality of the clustering results mined from the transformed data when the data are both sparse and dense?*

Our performance evaluation was carried out through the following steps:

- *Step 1*: we normalized the attribute values of the five real data sets using the z-score normalization. Normalization gives to all attributes the same weight.
- *Step 2*: we considered random projection based on two different approaches. First, the traditional way to compute random projection, by setting each entry of the random matrix R_1 to a value drawn from an i.i.d. $N(0, 1)$ distribution and then normalizing the columns to unit length. Second, we used the random matrix R_2 where each element r_{ij} is computed using Eq. (5). We refer to the former random

projection as RP_1 and the latter as RP_2. We repeated each experiment (for random projection) five times. In the next section, we present results by showing only the average value.

- *Step 3*: we computed the relative error that the distances in d–k space suffer from, on the average, by using the stress function given in Eq. (2). The stress function was computed for each data set.
- *Step 4*: we selected K-means to find the clusters in our performance evaluation. K-means is one of the best-known clustering algorithm and is scalable (Macqueen, 1967; Han et al., 2001).
- *Step 5*: we compared how closely each cluster in the transformed data set matches its corresponding cluster in the original data set. We expressed the quality of the generated clusters by computing the F-measure given in Eq. (10). Considering that K-means is not deterministic (due to its use of random seed selection), we repeated each experiment 10 times. We then computed the minimum, average, maximum, and standard deviation for each measured value of the F-measure. We present the results by showing only the average value.

We should point out that the steps described above were performed to evaluate the effectiveness of the DRBT when addressing PPC over centralized and vertically partitioned data.

5.3. *Evaluation approach*

When using random projection, a perfect reproduction of the Euclidean distances may not be the best possible result. The clusters in the transformed data sets should be equal to those in the original database. However, this is not always the case, and we have some potential problems after dimensionality reduction: (a) a noise data point ends up clustered; (b) a point from a cluster becomes a noise point; and (c) a point from a cluster migrates to a different cluster. In this research, we focus primarily on partitioning methods. In particular, we use K-means (Macqueen, 1967), one of the most used clustering algorithms. Since K-means is sensitive to noise points and clusters all the points in a data set, we have to deal with the third problem mentioned above (a point from a cluster migrates to a different cluster).

Our evaluation approach focuses on the overall quality of generated clusters after dimensionality reduction. We compare how closely each cluster in the transformed data matches its corresponding cluster in the original data set. To do so, we first identify the matching of clusters by computing the matrix of frequencies shown in Table 7. We refer to such a matrix as the clustering membership matrix (CMM), where the rows represent the clusters in the original data set, the columns represent the clusters in the transformed data set, and $freq_{i,j}$ is the number of points in cluster c_i that falls in cluster c'_j in the transformed data set.

Table 7 – The number of points in cluster c_i that falls in cluster c'_j in the transformed data set

	c'_1	c'_2	…	c'_k
c_1	$freq_{1,1}$	$freq_{1,2}$	…	$freq_{1,k}$
c_2	$freq_{2,1}$	$freq_{2,2}$	…	$freq_{2,k}$
⋮	⋮	⋮	⋱	⋮
c_k	$freq_{k,1}$	$freq_{k,2}$	…	$freq_{k,k}$

After computing the frequencies $freq_{i,j}$, we scan the clustering membership matrix calculating precision, recall, and F-measure for each cluster c'_j with respect to c_i in the original data set (Larsen and Aone, 1999). These formulas are given by the following equations:

$$\text{Precision } (P) = \frac{freq_{i,j}}{|c'_i|} \qquad (7)$$

$$\text{Recall } (R) = \frac{freq_{i,j}}{|c_i|} \qquad (8)$$

$$F\text{-measure } (F) = \frac{2PR}{(P+R)} \qquad (9)$$

where $|X|$ is the number of points in the cluster X.

For each cluster c_i, we first find a cluster c'_j that has the highest F-measure among all the c'_l, $1 \le l \le k$. Let $F(c_i)$ be the highest F-measure for cluster c_i, we denote the overall F-measure (OF) as the weighted average of $F(c_i)$, $1 \le i \le k$, as follows:

$$\text{OF} = \frac{\sum_{i=1}^{k} |c_i| F(c_i)}{\sum_{i=1}^{k} |c_i|} \qquad (10)$$

In the next sections, we present the performance evaluation results for clustering based on Eq. (10).

5.4. *Measuring the effectiveness of the DRBT over centralized data*

To measure the effectiveness of DRBT in PPC over centralized data, we started by computing the relative error that the distances in d–k space suffer from, on the average. To do so, we used the two random projection approaches (RP_1 and RP_2) mentioned in Step 3 of Section 5.2.

A word of notation: hereafter we denote the original dimension of a data set as d_o and reduced dimension of the transformed data set as d_r. This notation is to avoid confusion between the reduced dimension of a data set (k) and the number of clusters used as input of the algorithm K-means.

An important feature of the DRBT is its versatility to trade privacy, accuracy, and communication cost. The privacy preservation is assured because random projection is a non-invertible transformation, as discussed in Section 4.4. We here study the trade-off between accuracy and communication cost. The accuracy is represented by the error that the distances in d_o–d_r space suffer from, while the communication cost is represented by the number of dimensions that we reduce in the data sets. We selected two data sets: Pumsb and Chess with 74 and 37 dimensions, respectively. We reduced the dimensions of these data sets and computed the error. Fig. 4(a) shows the error produced by RP_1 and RP_2 on the data set Pumsb and Fig. 4(b) shows the error produced by RP_1 and RP_2 on the data set Chess. These results represent the average value of five trials. The error produced by RP_1 and RP_2 on the five data sets can be seen in Tables 8–12.

We observed that, in general, RP_2 yielded the best results in terms of the error produced on the data sets (the lower the

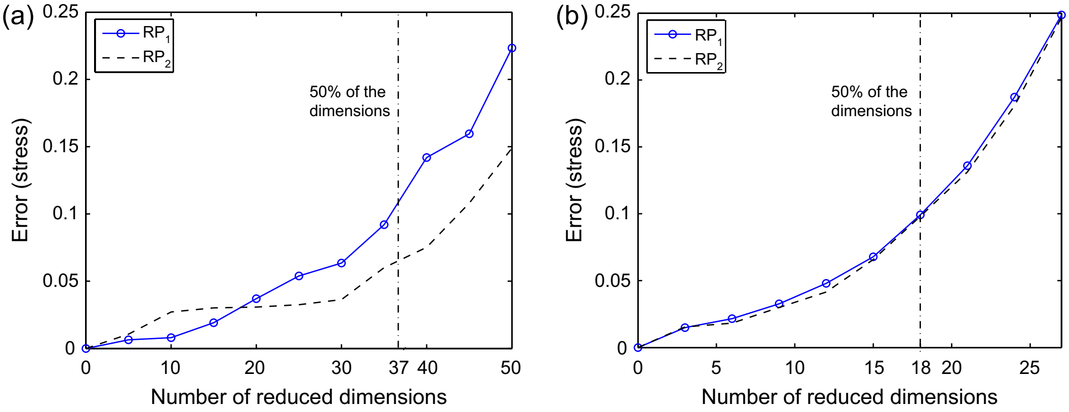


Fig. 4 – (a) The error produced on the data set Pumsb ($d_o = 74$). (b) The error produced in the data set Chess ($d_o = 37$).

better). In the data set Chess the difference between RP$_2$ and RP$_1$ was not significant. These results confirm the same findings in Bingham and Mannila (2001) and backup the theory of random projection (the choice of the random matrix) proposed in Achlioptas (2001). We noticed from the figures that the DRBT trades off accuracy (error) for communication cost (number of reduced dimensions) when the data are reduced up to 50% of the dimensions. In this case, the trade-off between the error and the communication cost is linear. However, reducing more than 50% of the dimensions, the communication cost is improved but the accuracy is compromised since the error produced on the data sets grows faster. Therefore, a data owner should consider carefully this trade-off before releasing some data for clustering.

After evaluating the error produced on the data sets, we used the algorithm *K*-means to find the clusters in the original and transformed data sets. We varied the number of clusters from 2 to 5 in the five data sets. Subsequently, we compared how closely each cluster in the transformed data set matches its corresponding cluster in the original data set by computing the *F*-measure given in Eq. (10).

Table 13 shows the results of the *F*-measure for the Accidents data set. We reduced the original 18 dimensions to 12. We repeated each experiment 10 times and computed the minimum, average, maximum, and standard deviation for each measured value of the *F*-measure. We simplify the results by showing only one data set (Accidents). The values of the *F*-measure for the other data sets can be found in Tables 14–17. Note that we computed the values of the *F*-measure only for the random projection RP$_2$ since its results were slightly better than those yielded by RP$_1$.

We noticed that the values of the *F*-measure for the Chess and Connect data sets (see Tables 14 and 17) were relatively low when compared with the results of the *F*-measure for the other data sets. The main reason is that the data points in these data sets are densely distributed. Thus, applying a partitioning clustering algorithm (e.g., *K*-means) to data sets of this nature increases the number of misclassified data points. On the other hand, when the attribute values of the objects are sparsely distributed, the clustering results are much better (see Tables 13, 15, and 16).

5.5. *Measuring the effectiveness of the DRBT over vertically partitioned data*

Now we move on to measure the effectiveness of DRBT to address PPC over vertically partitioned data. To do so, we split the Pumsb data set (74 dimensions) from 1 up to 4 parties (partitions) and fixed the number of dimensions to be reduced (38 dimensions). Table 18 shows the number of parties, the number of attributes per party, and the number of attributes in the merged data set which is subjected to clustering. Recall that in a vertically partitioned data approach, one of the parties will centralize the data before mining.

In this example, each partition with 37, 25, 24, 19, and 18 attributes was reduced to 19, 13, 12, 10, and 9 attributes, respectively. We applied the random projections RP$_1$ and RP$_2$ to each partition and then merged the partitions in one central repository. Subsequently, we computed the stress error on the merged data set and compared the error with the one produced on the original data set (without partitioning). Fig. 5 shows the error produced on the Pumsb data set in the vertically partitioned data approach. As we can see, the results yielded by RP$_2$ were again slightly better than those yielded by RP$_1$.

Table 8 – The error produced on the Chess data set ($d_o = 37$)

Chess	$d_r = 37$	$d_r = 34$	$d_r = 31$	$d_r = 28$	$d_r = 25$	$d_r = 22$	$d_r = 16$
RP$_1$	0.000	0.015	0.024	0.033	0.045	0.072	0.141
RP$_2$	0.000	0.014	0.019	0.032	0.041	0.067	0.131

Table 9 – The error produced on the Mushroom data set ($d_o = 23$)

Mushroom	$d_r = 23$	$d_r = 21$	$d_r = 19$	$d_r = 17$	$d_r = 15$	$d_r = 13$	$d_r = 9$
RP$_1$	0.000	0.020	0.031	0.035	0.048	0.078	0.155
RP$_2$	0.000	0.017	0.028	0.029	0.040	0.079	0.137

Table 10 – The error produced on the Pumsb data set ($d_o = 74$)

Pumsb	$d_r = 74$	$d_r = 69$	$d_r = 64$	$d_r = 59$	$d_r = 49$	$d_r = 39$	$d_r = 29$
RP_1	0.000	0.006	0.022	0.029	0.049	0.078	0.157
RP_2	0.000	0.007	0.030	0.030	0.032	0.060	0.108

Note that we reduced approximately 50% of the dimensions of the data set Pumsb and the trade-off between accuracy and communication cost is still efficient for PPC over vertically partitioned data.

We also evaluated the quality of clusters generated by mining the merged data set and comparing the clustering results with those mined from the original data set. To do so, we computed the *F*-measure for the merged data set in each scenario, i.e., from 1 up to 4 parties. We varied the number of clusters from 2 to 5. Table 19 shows values of the *F*-measure (average and standard deviation) for the Pumsb data set over vertically partitioned data. These values represent the average of 10 trials considering the random projection RP_2.

We notice from Table 19 that the results of the *F*-measure slightly decrease when we increase the number of parties in the scenario of PPC over vertically partitioned data. Despite this fact, the DRBT is still effective to address PPC over vertically partitioned data in preserving the quality of the clustering results as measured by *F*-measure.

5.6. *Discussion on the DRBT when addressing PPC*

The evaluation of the DRBT involves three important issues: security, communication cost, and quality of the clustering results. We discussed the issues of security in Section 4.4 based on Lemma 2, and the issues of communication cost and space requirements in Section 4.5. In this section, we have focused on the quality of the clustering results.

We have evaluated our proposed data transformation method (DRBT) to address PPC. We have learned some lessons from this evaluation, as follows:

- *The application domain of the DRBT*: we observed that the DRBT does not present acceptable clustering results in terms of accuracy when the data subjected to clustering are dense. Slightly changing the distances between data points by random projection results in misclassification, i.e., points will migrate from one cluster to another in the transformed data set. This problem is somehow understandable since partitioning clustering methods are not effective to find clusters in dense data. The Connect data set is one example which confirms this finding. On the other hand, our experiments demonstrated that the quality of the clustering results obtained from sparse data is promising.

Table 11 – The error produced on the Connect data set ($d_o = 43$)

Connect	$d_r = 43$	$d_r = 37$	$d_r = 31$	$d_r = 25$	$d_r = 19$	$d_r = 16$	$d_r = 13$
RP_1	0.000	0.016	0.037	0.063	0.141	0.159	0.219
RP_2	0.000	0.016	0.028	0.062	0.122	0.149	0.212

Table 12 – The error produced on the Accidents data set ($d_o = 18$)

Accidents	$d_r = 18$	$d_r = 16$	$d_r = 14$	$d_r = 12$	$d_r = 10$	$d_r = 8$	$d_r = 6$
RP_1	0.000	0.033	0.034	0.044	0.094	0.144	0.273
RP_2	0.000	0.018	0.023	0.036	0.057	0.108	0.209

- *The versatility of the DRBT*: using the DRBT, a data owner can tune the number of dimensions to be reduced in a data set trading privacy, accuracy, and communication costs before sharing the data set for clustering. Most importantly, the DRBT can be used to address PPC over centralized and vertically partitioned data.
- *The choice of the random matrix*: from the performance evaluation of the DRBT we noticed that the random projection RP_2 yielded the best results for the error produced on the data sets and the values of *F*-measure, in general. The random projection RP_2 is based on the random matrix proposed in Eq. (5).

6. Related work

Some effort has been made to address the problem of privacy preservation in data mining. The class of solutions for this problem has been restricted basically to *data partition*, *data modification*, *data restriction*, and *data ownership*.

6.1. *Data partitioning techniques*

Data partitioning techniques have been applied to some scenarios in which the databases available for mining are distributed across a number of sites, with each site only willing to share data mining results, not the source data. In these cases, the data are distributed either horizontally or vertically. In a horizontal partition, different entities are described with the same schema in all partitions, while in a vertical partition the attributes of the same entities are split across the partitions. The existing solutions can be classified into *Cryptography-Based Techniques* (Lindell and Pinkas, 2000; Kantarcioğlu and Clifton, 2002; Vaidya and Clifton, 2002, 2003) and *Generative-Based Techniques* (Meregu and Ghosh, 2003).

Table 13 – Average of the *F*-measure (10 trials) for the Accidents data set ($d_o = 18$, $d_r = 12$)

Data transformation	$k = 2$				$k = 3$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.931	0.952	0.941	0.014	0.903	0.921	0.912	0.009
Data transformation	$k = 4$				$k = 5$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.870	0.891	0.881	0.010	0.878	0.898	0.885	0.006

Table 14 – Average of *F*-measure (10 trials) for the Chess data set ($d_o = 37$, $d_r = 25$)

Data transformation	$k=2$				$k=3$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.529	0.873	0.805	0.143	0.592	0.752	0.735	0.050

Data transformation	$k=4$				$k=5$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.597	0.770	0.695	0.063	0.569	0.761	0.665	0.060

Table 16 – Average of *F*-measure (10 trials) for the Pumsb data set ($d_o = 74$, $d_r = 38$)

Data transformation	$k=2$				$k=3$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.611	0.994	0.909	0.140	0.735	0.991	0.965	0.081

Data transformation	$k=4$				$k=5$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.846	0.925	0.891	0.028	0.765	0.992	0.838	0.041

6.2. *Data modification techniques*

These techniques modify the original values of a database that needs to be shared, and in doing so, privacy preservation is ensured. The transformed database is made available for mining and must meet privacy requirements without losing the benefit of mining. In general, data modification techniques aim at finding an appropriate balance between privacy preservation and knowledge disclosure. Methods for data modification include *noise addition techniques* (Estivill-Castro et al., 1999; Agrawal and Srikant, 2000; Agrawal and Aggarwal, 2001; Evfimievski et al., 2002; Rizvi and Haritsa, 2002; Zang et al., 2004) and *space transformation techniques* (Oliveira and Zaïane, 2004).

The approach presented in this paper falls in the space transformation category. In this solution, the attributes of a database are reduced to a smaller number. The idea behind this data transformation is that by reducing the dimensionality of a database to a sufficiently small value, one can find a trade-off between privacy and accuracy. Once the dimensionality of a database is reduced, the released database preserves (or slightly modifies) the distances between data points. In addition, this solution protects individuals' privacy since the underlying data values of the objects subjected to clustering are completely different from the original ones.

6.3. *Data restriction techniques*

Data restriction techniques focus on limiting the access to mining results through either generalization or suppression of information (e.g., items in transactions, attributes in relations), or even by blocking the access to some patterns that are not supposed to be discovered. Such techniques can be divided into two groups: *Blocking-based techniques* (Johnsten and Raghavan, 1999, 2001; Saygin et al., 2001) and *Sanitization-based techniques* (Verykios et al., 2004; Dasseni et al., 2001; Oliveira and Zaïane, 2003; Oliveira et al., 2004; Iyengar, 2002).

6.4. *Data ownership techniques*

Data ownership techniques can be applied to two different scenarios: (1) to protect the ownership of data by people about whom the data were collected (Felty and Matwin, 2002). The idea behind this approach is that a data owner may prevent the data from being used for some purposes and allow them to be used for other purposes. To accomplish that, this solution is based on encoding permissions on the use of data as theorems about programs that process and mine the data. Theorem proving techniques are then used to guarantee that these programs comply with the permissions; and (2) to identify the entity that receives confidential data when such data are shared or exchanged (Mucsi-Nagy and Matwin, 2004). When sharing or exchanging confidential data, this approach ensures that no one can read confidential data except the receiver(s). It can be used in different scenarios, such as statistical or research purposes, data mining, and on-line business-to-business (B2B) interactions.

7. Conclusions

In this paper, we have showed analytically and experimentally that Privacy-Preserving Clustering (PPC) is possible to some extent. To support our claim, we introduced a new method to address PPC over centralized data and over

Table 15 – Average of *F*-measure (10 trials) for the Mushroom data set ($d_o = 23$, $d_r = 15$)

Data transformation	$k=2$				$k=3$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.972	0.975	0.974	0.001	0.689	0.960	0.781	0.105

Data transformation	$k=4$				$k=5$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.727	0.864	0.811	0.058	0.747	0.884	0.824	0.051

Table 17 – Average of *F*-measure (10 trials) for the Connect data set ($d_o = 43$, $d_r = 28$)

Data transformation	$k=2$				$k=3$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.596	0.863	0.734	0.066	0.486	0.863	0.623	0.103

Data transformation	$k=4$				$k=5$			
	Min	Max	Avg	Std	Min	Max	Avg	Std
RP_2	0.618	0.819	0.687	0.069	0.572	0.763	0.669	0.069

Table 18 – An example of partitioning for the Pumsb data set

No. of parties	No. of attributes per party	No. of attributes in the merged data set
1	1 Partition with 74 attributes	38
2	2 Partitions with 37 attributes	38
3	2 Partitions with 25 and 1 with 24 attributes	38
4	2 Partitions with 18 and 2 with 19 attributes	38

vertically partitioned data, called the Dimensionality Reduction-Based Transformation (DRBT). Our method was designed to support business collaboration considering privacy regulations, without losing the benefit of data analysis. The DRBT relies on the idea behind random projection to protect the underlying attribute values subjected to clustering. Random projection has recently emerged as a powerful method for dimensionality reduction. It preserves distances between data objects quite nicely, which is desirable in cluster analysis.

We evaluated the DRBT taking into account three important issues: security, communication cost, and accuracy (quality of the clustering results). Our experiments revealed that using DRBT, a data owner can meet privacy requirements without losing the benefit of clustering since the similarity between data points is preserved or marginally changed. From the performance evaluation, we suggested guidance on which scenario a data owner can achieve the best quality of the clustering when using the DRBT. In addition, we suggested guidance on the choice of the random matrix to obtain the best results in terms of the error produced on the data sets and the values of *F*-measure.

The highlights of the DRBT are as follows (a) it is independent of distance-based clustering algorithms; (b) it has a sound mathematical foundation; (c) it does not require CPU-intensive operations; and (d) it can be applied to address PPC over centralized data and PPC over vertically partitioned data.

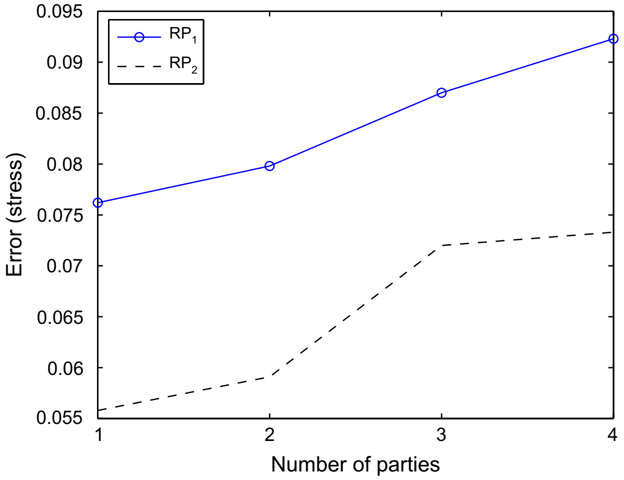


Fig. 5 – The error produced on the data set Pumsb over vertically partitioned data.

Table 19 – Average of the *F*-measure (10 trials) for the Pumsb data set over vertically partitioned data

No. of parties	$k=2$		$k=3$		$k=4$		$k=5$	
	Avg	Std	Avg	Std	Avg	Std	Avg	Std
1	0.909	0.140	0.965	0.081	0.891	0.028	0.838	0.041
2	0.904	0.117	0.931	0.101	0.894	0.059	0.840	0.047
3	0.874	0.168	0.887	0.095	0.873	0.081	0.801	0.073
4	0.802	0.155	0.812	0.117	0.866	0.088	0.831	0.078

REFERENCES

Achlioptas D. Database-friendly random projections. In: Proceedings of the 20th ACM symposium on principles of database systems. Santa Barbara, CA, USA; May 2001. p. 274–81.

Agrawal D, Aggarwal CC. On the design and quantification of privacy preserving data mining algorithms. In: Proceedings of ACM SIGMOD/PODS. Santa Barbara, CA; May 2001. p. 247–55.

Agrawal R, Srikant R. Privacy-preserving data mining. In: Proceedings of the 2000 ACM SIGMOD international conference on management of data. Dallas, Texas; May 2000. p. 439–50.

Auer JW. Linear algebra with applications. Scarborough, Ontario, Canada: Prentice-Hall Canada Inc.; 1991.

Berry M, Linoff G. Data mining techniques – for marketing, sales, and customer support. New York, USA: John Wiley and Sons; 1997.

Bingham E, Mannila H. Random projection in dimensionality reduction: applications to image and text data. In: Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining. San Francisco, CA, USA; 2001. p. 245–50.

Blake CL, Merz CJ. UCI repository of machine learning databases. Irvine: Department of Information and Computer Sciences, University of California; 1998.

Caetano TS. Graphical models and point set matching. PhD thesis, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil; July 2004.

Dasseni E, Verykios VS, Elmagarmid AK, Bertino E. Hiding association rules by using confidence and support. In: Proceedings of the fourth information hiding workshop. Pittsburg, PA; April 2001. p. 369–83.

Estivill-Castro V, Brankovic L. Data swapping: balancing privacy against precision in mining for logic rules. In: Proceedings of data warehousing and knowledge discovery DaWaK-99. Florence, Italy; August 1999. p. 389–98.

Evfimievski A, Srikant R, Agrawal R, Gehrke J. Privacy preserving mining of association rules. In: Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining. Edmonton, AB, Canada; July 2002. p. 217–28.

Faloutsos C, Lin K-I. FastMap: a fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. In: Proceedings of the 1995 ACM SIGMOD international conference on management of data. San Jose, CA, USA; June 1995. p. 163–74.

Felty AP, Matwin S. Privacy-oriented data mining by proof checking. In: Proceedings of the sixth European conference on principles of data mining and knowledge discovery (PKDD). Helsinki, Finland; August 2002. p. 138–49.

Fern XZ, Brodley CE. Random projection for high dimensional data clustering: a cluster ensemble approach. In: Proceedings of the 20th international conference on machine learning (ICML 2003). Washington DC, USA; August 2003.

Fukunaga K. Introduction to statistical pattern recognition. 2nd ed. Academic Press; 1990.

Han J, Kamber M. Data mining: concepts and techniques. San Francisco, CA: Morgan Kaufmann Publishers; 2001.

Iyengar VS. Transforming data to satisfy privacy constraints. In: Proceedings of the eighth ACM SIGKDDinternational conference on knowledge discovery and data mining. Edmonton, AB, Canada; July 2002. p. 279–88.

Jagadish HV. A retrieval technique for similar shapes. In: Proceedings of the 1991 ACM SIGMOD international conference on management of data. Denver, Colorado, USA; May 1991. p. 208–17.

Johnson WB, Lindenstrauss J. Extensions of Lipshitz mapping into Hilbert space. In: Proceedings of the conference in modern analysis and probability. Contemporary mathematics, vol. 26; 1984. p. 189–206.

Johnsten T, Raghavan VV. Impact of decision-region based classification mining algorithms on database security. In: Proceedings of 13th annual IFIP WG 11.3 working conference on database security. Seattle, USA; July 1999. p. 177–91.

Johnsten T, Raghavan VV. Security procedures for classification mining algorithms. In: Proceedings of 15th annual IFIP WG 11. 3 working conference on database and applications security. Niagara on the Lake, Ontario, Canada; July 2001. p. 293–309.

Kantarcioğlu M, Clifton C. Privacy-preserving distributed mining of association rules on horizontally partitioned data. In: Proceedings of the ACM SIGMOD workshop on research issues on data mining and knowledge discovery. Madison, Wisconsin; June 2002.

Kaski S. Dimensionality reduction by random mapping. In: Proceedings of the international joint conference on neural networks. Anchorage, Alaska; May 1999. p. 413–18.

Kruskal JB, Wish M. Multidimensional scaling. Beverly Hills, CA, USA: Sage Publications; 1978.

Larsen B, Aone C. Fast and effective text mining using linear-time document clustering. In: Proceedings of the fifth ACM SIGKDD international conference on knowledge discovery and data mining. San Diego, CA, USA; August 1999. p. 16–22.

Lindell Y, Pinkas B. Privacy preserving data mining, Crypto 2000. In: LNCS 1880. Santa Barbara, CA: Springer-Verlag; August 2000. p. 36–54.

Lo VSY. The true lift model – a novel data mining approach to response modeling in database marketing. SIGKDD Explorations December 2002;4(2):78–86.

Macqueen J. Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, vol. 1. Berkeley: University of California Press; 1967. p. 281–97.

Meregu S, Ghosh J. Privacy-preserving distributed clustering using generative models. In: Proceedings of the third IEEE international conference on data mining (ICDM'03). Melbourne, Florida, USA; November 2003. p. 211–18.

Mucsi-Nagy A, Matwin S. Digital fingerprinting for sharing of confidential data. In: Proceedings of the workshop on privacy and security issues in data Mining. Pisa, Italy; September 2004. p. 11–26.

Oliveira SRM, Zaïane OR. Protecting sensitive knowledge by data sanitization. In: Proceedings of the third IEEE international conference on data mining (ICDM'03). Melbourne, Florida, USA; November 2003. p. 613–16.

Oliveira SRM, Zaïane OR. Privacy-preserving clustering by object similarity-based representation and dimensionality reduction transformation. In: Proceedings of the workshop on privacy and security aspects of data mining (PSADM'04) in conjunction with the fourth IEEE international conference on data mining (ICDM'04). Brighton, UK; November 2004. p. 21–30.

Oliveira SRM, Zaïane OR, Saygin Y. Secure association rule sharing. In: Proceedings of the eighth Pacific–Asia conference on knowledge discovery and data mining (PAKDD'04). Sydney, Australia; May 2004. p. 74–85.

Papadimitriou CH, Tamaki H, Raghavan P, Vempala S. Latent semantic indexing: a probabilistic analysis. In: Proceedings of the 17th ACM symposium on principles of database systems. Seattle, WA, USA; June 1998. p. 159–68.

Rizvi SJ, Haritsa JR. Maintaining data privacy in association rule mining. In: Proceedings of the 28th international conference on very large data bases. Hong Kong, China; August 2002.

Samarati P. Protecting respondents' identities in microdata release. IEEE Transactions on Knowledge and Data Engineering 2001;13(6):1010–27.

Saygin Y, Verykios VS, Clifton C. Using unknowns to prevent discovery of association rules. SIGMOD Record December 2001;30(4):45–54.

Sweeney L. k-Anonymity: a model for protecting privacy. International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems 2002;10(5):557–70.

Vaidya J, Clifton C. Privacy preserving association rule mining in vertically partitioned data. In: Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining. Edmonton, AB, Canada; July 2002. p. 639–44.

Vaidya J, Clifton C. Privacy-preserving K-means clustering over vertically partitioned data. In: Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining. Washington, DC, USA; August 2003. p. 206–15.

Verykios VS, Elmagarmid AK, Bertino E, Saygin Y, Dasseni E. Association rule hiding. IEEE Transactions on Knowledge and Data Engineering 2004;16(4):434–47.

Young FW. Multidimensional scaling. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1987.

Zang N, Wang S, Zhao W. A new scheme on privacy preserving association rule mining. In: Proceedings of the 15th European conference on machine learning (ECML) and the eighth European conference on principles and practice of knowledge discovery in databases (PKDD). Pisa, Italy; September 2004.

Stanley R.M. Oliveira is a researcher at Brazilian Agricultural Research Corporation (Embrapa). He obtained his Ph.D. in Computer Science from the University of Alberta, Canada, in 2005. He obtained his B.S and M.Sc. degrees in Computer Science from the Federal University of Campina Grande, Paraiba, Brazil, in 1990 and 1996, respectively. He has published several papers in international and national venues, and a book chapter on privacy-preserving data mining. His main research interests include privacy-preserving data mining, database security, information hiding, and modeling and simulation.

Osmar R. Zaïane is an Associate Professor in Computing Science at the University of Alberta, Canada. Dr. Zaïane joined the University of Alberta in July of 1999. He obtained a Master's degree in Electronics at the University of Paris, France, in 1989 and a Master's degree in Computer Science at Laval University, Canada, in 1992. He obtained his Ph.D. from Simon Fraser University, Canada, in 1999 under the supervision of Dr. Jiawei Han. His Ph.D. thesis work focused on web mining and multimedia data mining. He has research interests in novel data mining algorithms, web mining, text mining, image mining, and information retrieval. He has published more than 80 papers in refereed international conferences and journals, and taught on all six continents. Osmar Zaïane was the co-chair of the ACM SIGKDD International Workshop on Multimedia Data Mining in 2000, 2001 and 2002 as well as co-chair of the ACM SIGKDD WebKDD workshop in 2002, 2003 and 2005. He is the Program co-chair for the IEEE International Conference on Data Mining 2007 and General co-chair for the conference on Advanced Data Mining Applications 2007. Osmar Zaïane is the ACM SIGKDD Explorations Associate Editor and Associate Editor of the International Journal of Internet Technology and Secured Transactions.

ELSEVIER

Computers & Security

Simple three-party key exchange protocol

Rongxing Lu, Zhenfu Cao[*]

Department of Computer Science and Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, People's Republic of China

ARTICLE INFO

Article history:
Received 13 July 2005
Accepted 3 August 2006

Keywords:
Protocol
Key exchange
Three-party key exchange
Password based

ABSTRACT

Three-party authenticated key exchange protocol is an important cryptographic technique in the secure communication areas, by which two clients, each shares a human-memorable password with a trusted server, can agree a secure session key. Over the past years, many three-party authenticated key exchange protocols have been proposed. However, to our best knowledge, not all of them can meet the requirements of security and efficiency simultaneously. Therefore, in this paper, we would like to propose a new simple three-party password based authenticated key exchange protocol. Compared with other existing protocols, our proposed protocol does not require any server's public key, but can resist against various known attacks. Therefore, we believe it is suitable for some practical scenarios.

1. Introduction

In the secure communication areas, key exchange protocol is one of the most important cryptographic mechanisms, by which a pair of users that communicate over a public unreliable channel can generate a secure session key to guarantee the later communications' privacy and data integrity. In the seminal paper of Diffie and Hellman (1976), they proposed the first practical key exchange protocol. However, the Diffie–Hellman protocol does not provide the authentication mechanism, and therefore easily suffers from the "man-in-the-middle" attack. To resolve this issue, over the past years, a bulk of key exchange protocols with authentication function have been developed (Blake-Wilson et al., 1997; Law et al., 2003; Zhang et al., 2002; Boyd et al., 2004; Bellovin and Merrit, 1992; Abdalla and Pointcheval, 2005), amongst which the password based authenticated key exchange protocol receives much interest (Bellovin and Merrit, 1992; Abdalla and Pointcheval, 2005).

Because the password based authenticated key exchange protocols require users only to remember a human-memorable (low-entropy) password, it is rather simple and efficient. However, just as the chosen passwords are of low-entropy, it is not trivial to protect the password information against the password guessing attacks (dictionary attacks). Therefore, since Bellovin and Merrit (1992) proposed the first password based authenticated key exchange (PAKE) protocol, it has been widely studied, and many excellent protocols have been developed.

In recent years, besides the two-party PAKE protocols, many researchers also began to consider the three-party PAKE protocols (Steiner et al., 1995; Ding and Horster, 1995; Lin et al., 2000, 2001; Chang and Chang, 2004; Lee et al., 2004, 2005; Sun et al., 2005). In a three-party PAKE protocol, each client first shares a human-memorable password with a trusted server, and then when two clients want to agree a session key, they resort to the trusted server for authenticating each other. In 1995, Steiner et al. proposed a three-party PAKE (3PAKE) protocol.

[*] *Corresponding author*. Tel.: +86 21 3420 4642.
E-mail addresses: rxlu.cn@gmail.com (R. Lu), cao-zf@cs.sjtu.edu.cn (Z. Cao).

Although a trusted server can help the two clients to authenticate each other and agree a common session key, yet their 3PAKE protocol still suffers from password guessing attacks. Ding and Horster (1995) showed that Steiner et al.'s 3PAKE protocol is vulnerable to undetectable on-line password guessing attacks. And in 2000, Lin et al. also pointed out Steiner et al.'s 3PAKE protocol suffers from off-line password guessing attacks, and then presented an improved version. Since Lin et al.'s (2000) improved 3PAKE protocol applied the server's public key; it can resist the password guessing attacks. However, employing the server's public key also puts a burden on the clients, because they have to verify it beforehand. Therefore, to reduce the clients' burdens, in 2001, Lin et al. presented another improved protocol without server's public key. However, this improved protocol needs two more rounds than the improved protocol in Law et al. (2003). More recently, Chang and Chang (2004) employed super-poly-one trapdoor functions to design a novel three-party encrypted key exchange protocol, and Lee et al. (2004) presented an enhanced three-party encrypted key exchange protocol without server public keys and its round efficient version. On the other hand, another two verifier-based protocols have also been proposed recently by Sun et al. (2005) and Lee et al. (2005).

In this paper, motivated by Abdalla and Pointcheval's (2005) simple password based encrypted key exchange protocol (SPAKE), we would like to propose a new simple three-party key exchange protocol (S-3PAKE). Since our proposed protocol does not require any server's public keys, it seems very simple and efficient, and can be used in many practical scenarios.

The rest of this paper is organized as follows. In Section 2, we first review Abdalla and Pointcheval's SPAKE protocol. Then, we present our simple three-party key exchange S-3PAKE protocol in Section 3. And then we analyze the protocol's security and efficiency in Sections 4 and 5, respectively. Finally, we draw our conclusions in Section 6.

2. Simple password based encrypted key exchange protocol

Abdalla and Pointcheval (2005) suggested a new variation of the computational Diffie–Hellman assumption, called chosen-based computational Diffie–Hellman assumption, and presented an elegant simple password based encrypted key exchange protocol (SPAKE) based on this new assumption. In this section, we first review such a simple SPAKE protocol.

Computational Diffie–Hellman (CDH): let G be a finite cyclic group, and let g be the generator of prime order p in G. The CDH problem is stated as "given g^x, g^y, where $x, y \in Z_p$, to compute $\mathrm{CDH}(g^x, g^y) = g^{xy}$".

Chosen-based computational Diffie–Hellman (CCDH) Abdalla and Pointcheval (2005): the CCDH problem is a variation of the CDH problem. It mainly considers an adversary that is given three random elements M, N and X in G and whose goal is to find a triple of values (Y, u, v) such that $u = \mathrm{CDH}(X, Y)$ and $v = \mathrm{CDH}(X/M, Y/N)$. The idea behind this assumption is that the adversary may be able to successfully compute either u or v, but not both. Abdalla and Pointcheval (2005) proved that solving CCDH problem is equivalent to solving the underlying CDH problem in G.

SPAKE is a two-party key exchange protocol, which can be regarded as a variation of Bellovin and Merrit's (1992) protocol. In the SPAKE protocol, the system parameters are (G, g, p, M, N, H), where $M, N \in G$ and H is a one-way hash function. Now, assume that users A and B share a low-entropy password pw and want to agree a common session key, the SPAKE protocol, as shown in Fig. 1, will run as follows.

Step 1:
A1: user A first chooses a random number $x \in Z_p$ and computes $X \leftarrow g^x$, $X^* \leftarrow X \cdot M^{pw}$, then sends X^* to user B.
B1: similarly, user B also chooses a random number $y \in Z_p$ and computes $Y \leftarrow g^y$, then sends $Y^* \leftarrow Y \cdot N^{pw}$ to user A.
Step 2:
A2: upon receiving Y^*, user A computes $K_A \leftarrow (Y^*/N^{pw})^x$, then computes the session key $SK_A \leftarrow H(A, B, X^*, Y^*, K_A)$.
B2: similarly, when user B receives X^*, he computes $K_B \leftarrow (X^*/M^{pw})^y$, and computes the session key $SK_B \leftarrow H(A, B, X^*, Y^*, K_B)$.

According to the fact that $K_A = K_B = g^{xy}$, the SPAKE protocol's correctness is obviously satisfied.

Based upon the CCDH assumption, Abdalla and Pointcheval (2005) also have proved that the SPAKE protocol is secure in the random oracle model. Thus, the SPAKE protocol is not only efficient but also secure. In the next section, we will present our simple three-party password based key exchange protocol (S-3PAKE) based on the SPAKE protocol.

3. Simple three-party key exchange protocol

In this section, we present our simple three-party password based key exchange protocol (S-3PAKE). To illustrate the protocol clear, we first introduce some notations in Section 3.1, and then describe the concrete protocol in Section 3.2.

3.1. Notations

Here some notations used in the proposed protocol are first listed as follows:

(G, g, p): a finite cyclic group G generated by an element g of prime order p.

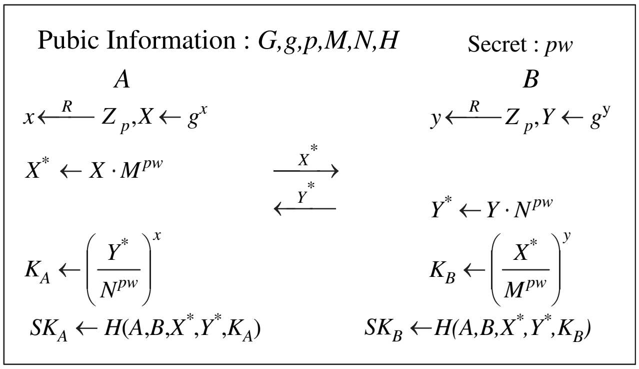

Pubic Information : G,g,p,M,N,H
Secret : pw
A
B
$x \xleftarrow{R} Z_p, X \leftarrow g^x$
$y \xleftarrow{R} Z_p, Y \leftarrow g^y$
$X^* \leftarrow X \cdot M^{pw}$
X^*
Y^*
$Y^* \leftarrow Y \cdot N^{pw}$
$K_A \leftarrow \left(\frac{Y^*}{N^{pw}}\right)^x$
$K_B \leftarrow \left(\frac{X^*}{M^{pw}}\right)^y$
$SK_A \leftarrow H(A,B,X^*,Y^*,K_A)$
$SK_B \leftarrow H(A,B,X^*,Y^*,K_B)$

Fig. 1 – Simple password based encrypted key exchange protocol.

M, N: two elements in G.
S: a trusted server.
A, B: two clients.
pw1: the password shared between A and S.
pw2: the password shared between B and S.
H, H': two secure one-way hash functions.

3.2. The proposed protocol

In this system, assume that two clients A and B wish to agree a common session key. However, as they do not hold any shared information in advance, they cannot directly authenticate each other and have to resort to the trusted server S for a session key agreement. The detailed steps of the S-3PAKE protocol, as shown in Fig. 2, are described as follows:

Step 1:
A1: A chooses a random number $x \in Z_q$ and computes $X \leftarrow g^x \cdot M^{pw1}$, then sends $A \| X$ to B.
B1: B also chooses a random number $y \in Z_p$ and computes $Y \leftarrow g^y \cdot N^{pw2}$, then sends $A\|X\|B\|Y$ to S.
Step 2:
S2: upon receiving $A\|X\|B\|Y$, the server S first uses the passwords pw1 and pw2 to compute $g^x \leftarrow X/M^{pw1}$ and $g^y \leftarrow Y/N^{pw2}$, respectively. Then, she chooses another random number $z \in Z_p$ and computes $g^{xz} \leftarrow (g^x)^z$, $g^{yz} \leftarrow (g^y)^z$. Finally, she sends $X'\|Y'$ to B, where $X' \leftarrow g^{yz} \cdot H(A,S,g^x)^{pw1}$ and $Y' \leftarrow g^{xz} \cdot H(A,S,g^y)^{pw2}$.
B2: when B receives $X'\|Y'$, he uses the password pw2 to compute $g^{xz} \leftarrow Y'/H(B,S,g^y)^{pw2}$, and uses the random number y to compute $g^{xyz} \leftarrow g^{(xz)y}$. At last, he forwards $X'\|\alpha$ to A, where $\alpha \leftarrow H(A,B,g^{xyz})$.
Step 3:
A3: after A receives $X'\|\alpha$, she first computes $g^{yz} \leftarrow X'/H(B,S,g^x)^{pw1}$ and $g^{xyz} \leftarrow g^{(yz)x}$. Then, she checks whether $\alpha = H(A,B,g^{xyz})$ holds or not. If it does not hold, A terminates the protocol. Otherwise, she is convinced that g^{xyz} is valid. And in this case, she can compute the session key $SK_A \leftarrow H'(A,B,g^{xyz})$ and returns $\beta \leftarrow H(B,A,g^{xyz})$ to B for validation.
B3: upon receiving β, B checks whether $\beta = H(B,A,g^{xyz})$ holds or not. If it does hold, B can compute the session key $SK_B \leftarrow H'(A,B,g^{xyz})$. Otherwise, he terminates the protocol.

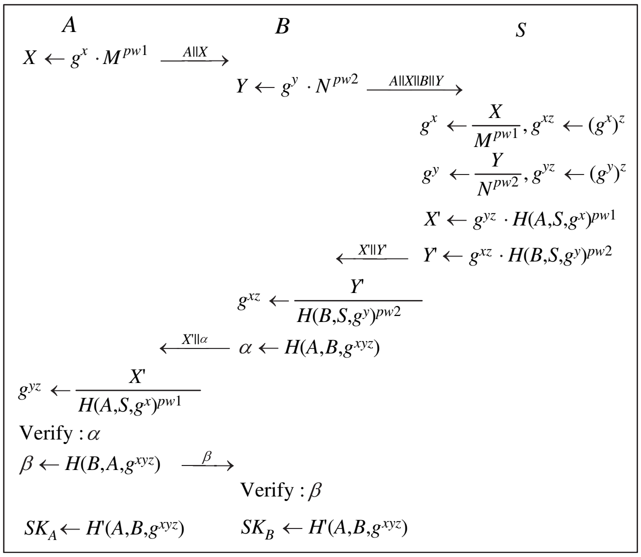

A
B
S
$X \leftarrow g^x \cdot M^{pw1}$
A||X
$Y \leftarrow g^y \cdot N^{pw2}$
A||X||B||Y
$g^x \leftarrow \frac{X}{M^{pw1}}, g^{xz} \leftarrow (g^x)^z$
$g^y \leftarrow \frac{Y}{N^{pw2}}, g^{yz} \leftarrow (g^y)^z$
$X' \leftarrow g^{yz} \cdot H(A,S,g^x)^{pw1}$
X'||Y'
$Y' \leftarrow g^{xz} \cdot H(B,S,g^y)^{pw2}$
$g^{xz} \leftarrow \frac{Y'}{H(B,S,g^y)^{pw2}}$
X'||α
$\alpha \leftarrow H(A,B,g^{xyz})$
$g^{yz} \leftarrow \frac{X'}{H(A,S,g^x)^{pw1}}$
Verify : α
$\beta \leftarrow H(B,A,g^{xyz})$
β
Verify : β
$SK_A \leftarrow H'(A,B,g^{xyz})$
$SK_B \leftarrow H'(A,B,g^{xyz})$

Fig. 2 – Simple three-party key exchange protocol.

Correctness: from the above steps, if A, B, S all follow the proposed protocol, and both α and β are accepted, then A and B can derive the same session key $SK_A = SK_B = H'(A,B,g^{xyz})$. Thus, the correctness follows.

4. Security discussions

In this section, we will analyze that our proposed S-3PAKE protocol is secure and can work correctly. Since our protocol is derived from the SPAKE protocol (Abdalla and Pointcheval, 2005), and the SPAKE protocol is provably secure in the random oracle model. We believe our protocol may inherit some security properties. Therefore, here we mainly discuss whether our proposed S-3PAKE protocol can resist various known attacks.

Trivial attack: an attacker may directly try to compute the passwords and/or the session key from the transmitted transcripts (X, Y, X', Y', α, β). However, due to the difficulties of the discrete logarithm problem and CCDH problem and the one-wayness of hash function H, the trivial attack is useless to our proposed S-3PAKE protocol.

On-line guessing attack: an attacker may try to guess and obtain A's password. In order to do that, the attacker first guesses a password pw1* and sends $X^* \leftarrow g^x \cdot M^{pw1^*}$ to B. After some executions of S and B, the attacker will receive $X'\|\alpha$. If the guessing is wrong, the attacker obviously cannot verify α or produce a valid β, thus B can detect the user is a forger immediately. Similarly, the attacker also cannot guess B's password. Hence the on-line guessing attack is invalid to our proposed S-3PAKE protocol.

Off-line guessing attack: in an off-line guessing attack, an attacker guesses a password and confirms his guess off-line. However, like the SPAKE protocol (Abdalla and Pointcheval, 2005), there is also no useful information to help verify the correctness of the guessed passwords in our proposed S-3PAKE protocol, as H is one-way hash function and x, y, z are all random numbers. Therefore, our proposed S-3PAKE protocol can resist the off-line guessing attack.

Replay attack: in a replay attack, an attacker may want to pretend to be A by replaying X to B. However, as he does not know x, and y, z are randomly chosen in each session, the attacker has no ability to create a correct β and produce a valid session key. Similarly, the attacker also cannot pretend to be B. Hence, the replay attack is of no use in our proposed S-3PAKE protocol.

Perfect forward secrecy: the proposed S-3PAKE protocol also has the perfect forward secrecy. Even if an attacker has compromised two passwords pw1 and pw2, he still cannot learn any previous session keys, because the corresponding ephemeral parameters x, y, z in each session are randomly chosen and of independence.

Known-key security: in our proposed S-3PAKE protocol, as the ephemeral parameters x, y, z in each session key are

Table 1 – Performance comparison of round efficient 3PAKE and our proposed S-3PAKE

	Round efficient 3PAKE (Lee et al., 2004)			Our proposed S-3PAKE		
	A	*B*	*S*	*A*	*B*	*S*
Modular exponentiation	3	3	4	3(1)	3(1)	4(2)
Hash/pseudo-random	6	6	4	3	3	2
System en(decryption)	1	1	2	0	0	0
Random number	1	1	2	1	1	1
Round number		4			3	

random and independent of other session keys. Therefore, the knowledge of previous session keys does not help to derive a new session key. Hence, known-key security is satisfied in our proposed protocol.

In the end, from what has been analyzed above, we are fully convinced that our proposed S-3PAKE protocol is secure and can work correctly.

5. Performance analyses

In this section, we will show that our proposed S-3PAKE protocol is also an efficient one. As Lee et al.'s (2004) round efficient 3PAKE protocol is more efficient than other protocols (Lin et al., 2000, 2001), here we use Table 1 to show the performance comparisons of the round efficient 3PAKE protocol (Lee et al., 2004) and our proposed S-3PAKE protocol. The comparison factors include the number of rounds, random numbers, hash/pseudo-random functions, exponentiations and symmetric en/decryption.

As shown in Table 1, if we consider some information, such as M^{pw1} and N^{pw2}, can be pre-computed in the protocol, then our proposed S-3PAKE protocol is obviously more efficient than Lee et al.'s (2004) round efficient 3PAKE protocol. Therefore, we believe our proposed S-3PAKE protocol is really a simple and efficient three-party key exchange protocol, and can be applied in practice.

6. Conclusions

In this paper, we first reviewed a simple password based encrypted key exchange (SPAKE) protocol due to Abdalla and Pointcheval (2005), and then extended it to a simple three-party password based key exchange (S-3PAKE) protocol without server public key. By analysis, our proposed S-3PAKE protocol is not only secure against various known attacks, but also highly efficient. Therefore, we believe it is particularly useful in some practical scenarios.

Acknowledgements

This work is supported in part by the National Natural Science Foundation of China for Distinguished Young Scholars under Grant Nos. 60225007 and 60572155, the Science and Technology Research Project of Shanghai under Grant No. 04DZ07067, and the Special Research Funds of Huawei.

REFERENCES

Abdalla M, Pointcheval D. Simple password-based encrypted key exchange protocols, Topics in cryptology – CT-RSA 2005. In: LNCS. Springer-Verlag; 2005. p. 191–208.

Bellovin SM, Merrit M. Encrypted key exchange: password-based protocols secure against dictionary attacks. In: Proceedings of IEEE symposium on research in security and privacy. IEEE Computer Society Press; May 1992. p. 72–84.

Blake-Wilson S, Johnson D, Menezes A. Key agreement protocols and their security analysis. In: Proceedings of sixth IMA international conference on cryptography and coding. LNCS 1355. Springer-Verlag; 1997. p. 30–45.

Boyd C, Mao W, Paterson K. Key agreement using statically keyed authenticators, Applied cryptography and network security – ACNS'2004. In: LNCS 3089. Springer-Verlag; 2004. p. 248–62.

Chang CC, Chang YF. A novel three-party encrypted key exchange protocol. Computer Standards and Interfaces 2004;26(5):471–6.

Diffie W, Hellman M. New directions in cryptography. IEEE Transactions on Information Theory, IT November 1976;22(6):644–54.

Ding Y, Horster P. Undetectable on-line password guessing attacks. ACM Operating Systems Review 1995;29(4):77–86.

Law L, Menezes A, Qu M, Solinas J, Vanstone S. An efficient protocol for authenticated key agreement. Designs, Codes and Cryptography March 2003;28(2):119–34.

Lee TF, Hwang T, Lin CL. Enhanced three-party encrypted key exchange without server public keys. Computers and Security 2004;23(7):571–7.

Lee SW, Kim HS, Yoo KY. Efficient verifier-based key agreement protocol for three parties without server's public key. Applied Mathematics and Computation 2005;167(2):996–1003.

Lin CL, Sun HM, Hwang T. Three party-encrypted key exchange: attacks and a solution. ACM Operating Systems Review 2000; 34(4):12–20.

Lin CL, Sun HM, Steiner M, Hwang T. Three-party encrypted key exchange without server public-keys. IEEE Communication Letters 2001;5(12):497–9.

Steiner M, Tsudik G, Waidner M. Refinement and extension of encrypted key exchange. ACM Operating Systems Review 1995;29(3):22–30.

Sun HM, Chen BC, Hwang T. Secure key agreement protocols for three-party against guessing attacks. The Journal of Systems and Software 2005;75:63–8.

Zhang F, Liu S, Kim K. ID-based one round authenticated tripartite key agreement protocol with pairings, Cryptology ePrint archive, report 2002/122. Available from: http://eprint.iarc.org/2002/122; 2002.

Rongxing Lu received his B.S. and M.S. degrees in computer science from Tongji University in 2000 and 2003, respectively. Currently, he is a doctoral candidate in the Department of Computer and Engineering, Shanghai Jiao Tong University. His research interests lie in cryptography and network security.

Zhenfu Cao is the professor and the doctoral supervisor of Computer Software and Theory at Department of Computer Science of Shanghai Jiao Tong University. His main research areas are number theory and modern cryptography, theory and technology of information security, etc. He is the gainer of Ying-Tung Fok Young Teacher Award (1989), the First 10 Outstanding Youth in Harbin (1996), Best PhD thesis award in Harbin Institute of Technology (2001), and the National Outstanding Youth Fund in 2002.

Refereed papers — Guide for Authors

Scope

Computers & Security is the most comprehensive, authoritative survey of the key issues in computer security today. It aims to satisfy the needs of managers and experts involved in the computer security field by providing a combination of leading edge research developments, innovations and sound practical management advice for computer security professionals worldwide. *Computers & Security* provides detailed information to the professional involved with computer security, audit, control and data integrity in all sectors — industry, commerce and academia.

Submissions

Original submissions on all computer security topics are welcomed, especially those of practical benefit to the computer security practitioner.
From 1 April 2006, submissions with cryptology theory as their primary subject matter will no longer be accepted by *Computers & Security* as anything other than invited contributions. Authors submitting papers that feature cryptologic results as an important supporting feature should ensure that the paper, as a whole, is of importance to the advanced security practitioner or researcher, and ensure that the paper advances the overall field in a significant manner. Authors who submit purely theoretical papers on cryptology may be advised to resubmit them to a more appropriate journal; the Editorial Board reserves the right to reject such papers without the full reviewing process. Cryptography papers submitted before this date will be subject to the usual reviewing process, should the paper pass the pre-review process which has been in place since 2004.

All contributions should be in English and, since the readership of the journal is international, authors are reminded that simple, concise sentences are our preferred style. It is also suggested that papers are spellchecked and, if necessary, proofread by a native English speaker in order to avoid grammatical errors. All technical terms that may not be clear to the reader should be clearly explained.

Copyright is retained by the Publisher. Submission of an article implies that the paper has not been published previously; that it is not under consideration for publication elsewhere; that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out; and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher.

All papers will be submitted to expert referees from the editorial board for review. The usual size of a paper is 5000 to 10 000 words. Please contact n.dudley@elsevier.com if further clarification is needed.

Please ensure that the title contains all the authors' names, affiliations, and their full mailing addresses. These should be followed by a brief abstract and a list of five to 10 keywords. Please supply figures and tables separately. Figures should be high resolution computer-generated graphics, clearly printed black and white line drawings, or black and white glossy photographs. All illustrations should be large enough to withstand 50% reduction and still be easily readable. Try to incorporate all material into the text, and avoid footnotes wherever possible. Any measurements must be in SI (Système International) units.

References should be consecutively numbered throughout the text and then listed in full at the end of the paper.

Accepted papers

If the paper is accepted, or accepted subject to revision, the authors are requested to send a digital copy of the final version of the paper. Please supply the digital file as either a Microsoft Word file or as a LaTeX file, together with an Adobe Acrobat PDF. Please supply figures and tables as separate files. We will also need a short (100 words) biographical sketch of each author.

A copy of the relevant journal issue will be supplied free of charge to the main author. Twenty five reprints can be provided on request. Further reprints (minimum order of 100) can be supplied at a reasonable cost if the request is received before the issue goes to press.

Papers or abstracts for discussion should be submitted to:

Submission for all types of manuscripts to *Computers & Security* now proceeds totally online.

Via the Elsevier Editorial System Website for this journal at http://ees.elsevier.com/cose, you will be guided stepwise through the creation and uploading of the various files. When submitting a manuscript to Elsevier

Editorial System, authors need to provide an electronic version of their manuscript. For this purpose only original source files are allowed, so no PDF files. Authors should select a category designation for their manuscripts (network intrusion, security management etc). Authors may send queries concerning the submission process, manuscript status, or journal procedures to Author Support at authorsuppor@elsevier.com. Once the uploading is completed, the system generates an electronic (PDF) proof, which is then used for reviewing. All correspondence, including the editor's decision and request for revisions, will be by e-mail.

If online submission is not possible, manuscripts may be submitted by sending the source files via email attachment (please note that this is no longer the preferred way of submission and could cause a considerable delay in publication of the article) to c.dijk@elsevier.com

ELSEVIER

Computers
&
Security

www.elsevier.com/locate/cose

EVENTS

RSA CONFERENCE USA 2007
5–9 February 2007
San Francisco, USA
www.rsaconference.com

22ND ANNUAL ACM SYMPOSIUM ON APPLIED COMPUTING
11–15 March 2007
Seoul, Korea
http://comp.uark.edu/%7Ebpanda/sac-cf.htm

ARCHITECTURE OF COMPUTING SYSTEMS
12–15 March 2007
Zurich, Switzerland
http://arcs07.ethz.ch

ISACA EUROCACS
18–21 March 2007
Vienna, Austria
www.isaca.org

CYBERCRIME SUMMIT
19–23 March 2007
Atlanta, Georgia, USA
www.southeastcybercrimesummit.com/index.htm

E-CRIME CONGRESS
27–28 March 2007
London, UK
www.e-crimecongress.org

5TH INTL. SYMPOSIUM ON MODELING AND OPTIMIZATION IN MOBILE, AD HOC, AND WIRELESS NETWORKS
16–20 April 2007
Limassol, Cyprus
www.wiopt.org/

CSRC 6TH ANNUAL PKI R&D CONFERENCE
17–19 April
Gathersburg, MD, USA
http://csrc.nist.gov/events

INFOSECURITY EUROPE
24–26 April 2007
London, UK
www.infosec.co.uk

NEW TECHNOLOGIES, MOBILITY AND SECURITY 2007
30 April –5 May 2007
Beirut, Lebanon
www.ntms2007.org

16TH INTERNATIONAL WORLD WIDE WEB CONFERENCE
8–12 May 2007
Banff, Canada
http://www2007.org